(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119072486 A
(43)申请公布日 2024.12.03

(21)申请号 202380031870.2
(22)申请日 2023.02.07
(30)优先权数据
63/307,742 2022.02.08 US
63/339,574 2022.05.09 US
(85)PCT国际申请进入国家阶段日
2024.09.29
(86)PCT国际申请的申请数据
PCT/US2023/062101 2023.02.07
(87)PCT国际申请的公布数据
WO2023/154693 EN 2023.08.17
(71)申请人 沃雅戈治疗公司
地址 美国马萨诸塞州
(72)发明人 M·E·农嫩马赫 A·Z·任 D·莫拉 王伟
(74)专利代理机构 隆天知识产权代理有限公司 72003
专利代理师 付文川
(51)Int.Cl.
C07K 14/015(2006.01)
C12N 15/35(2006.01)
C07K 7/04(2006.01)
C12N 15/863(2006.01)
A61K 48/00(2006.01)
C40B 40/08(2006.01)
C12N 15/10(2006.01)

权利要求书12页 说明书231页
序列表（电子公布） 附图4页

(54)发明名称

AAV衣壳变体及其用途

(57)摘要

本公开涉及用于制备、使用和/或配制腺相关病毒衣壳蛋白变体的组合物和方法。

CN 119072486 A

1.一种AAV衣壳变体,其包含具有下式的氨基酸序列:

[N2]-[N3],其中:

(i)[N2]包含位置X1、X2、X3、X4和X5,其中:

(a)位置X1为Y、N或C;

(b)位置X2为P、K、T或Q;

(c)位置X3为A或P;

(d)位置X4为E、S或A;且

(e)位置X5为V、L或E;和

(ii)[N3]包含VQK、EQK、VKK、VHK、VQQ或LQK的氨基酸序列;并且

其中所述AAV衣壳变体包含SEQ ID NO:739的氨基酸序列,或与其至少95%同一的氨基酸序列。

2.如权利要求1所述的AAV衣壳变体,其中:

(a)[N2]包含YP、NK、YT、YQ、NP、CP、TH、AE、PS、AA、AS、PA、PP、KA、TA、QA、TP、HA、EV、SL、EE、AV或SH;

(b)[N2]包含YPA、YPP、NKA、YTA、YQA、YTP、NPA、CPA、THA、PAE、PPS、KAE、TAE、QAE、TPS、PAA、HAS、AEV、PSL、AEE或AAV;

(c)[N2]包含YPAE(SEQ ID NO:21)、YPPS(SEQ ID NO:22)、NKAE(SEQ ID NO:23)、YTAE(SEQ ID NO:24)、YQAE(SEQ ID NO:25)、YTPS(SEQ ID NO:26)、YPAA(SEQ ID NO:27)、NPAE(SEQ ID NO:28)、CPAE(SEQ ID NO:29)、THAS(SEQ ID NO:30)、PAEV(SEQ ID NO:17)、PPSL(SEQ ID NO:31)、KAEV(SEQ ID NO:32)、TAEV(SEQ ID NO:16)、PAEE(SEQ ID NO:18)、QAEV(SEQ ID NO:15)、TPSL(SEQ ID NO:33)、PAAV(SEQ ID NO:34)或QAEE(SEQ ID NO:35);和/或

(d)[N2]包含YPAEV(SEQ ID NO:1)、YPPSL(SEQ ID NO:2)、NKAEV(SEQ ID NO:3)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、YQAEV(SEQ ID NO:6)、YTPSL(SEQ ID NO:7)、YPAAV(SEQ ID NO:8)、NPAEV(SEQ ID NO:9)、CPAEV(SEQ ID NO:10)或YQAEE(SEQ ID NO:11)。

3.如权利要求1或2所述的AAV衣壳变体,其中[N2]-[N3]包含:

(i)AEVVQK(SEQ ID NO:36)、PSLVQK(SEQ ID NO:37)、AEVEQK(SEQ ID NO:38)、AEEVQK(SEQ ID NO:39)、PSLEQK(SEQ ID NO:40)、PSLVKK(SEQ ID NO:41)、AEVVKK(SEQ ID NO:42)、AEVVHK(SEQ ID NO:43)、AAVVQK(SEQ ID NO:44)、AEVVQQ(SEQ ID NO:45)或AEVLQK(SEQ ID NO:46)的氨基酸序列

(ii)PAEVVQK(SEQ ID NO:20)、PPSLVQK(SEQ ID NO:47)、KAEVVQK(SEQ ID NO:48)、TAEVVQK(SEQ ID NO:49)、PAEVEQK(SEQ ID NO:50)、PAEEVQK(SEQ ID NO:51)、QAEVVQK(SEQ ID NO:52)、TPSLVQK(SEQ ID NO:53)、PPSLEQK(SEQ ID NO:54)、PPSLVKK(SEQ ID NO:55)、PAEVVKK(SEQ ID NO:56)、PAEVVHK(SEQ ID NO:57)、PAAVVQK(SEQ ID NO:58)、PAEVVQQ(SEQ ID NO:59)、TAEVVKK(SEQ ID NO:60)、PAEVLQK(SEQ ID NO:61)或QAEEVQK(SEQ ID NO:62)的氨基酸序列。

4.如权利要求1至3中任一项所述的AAV衣壳变体,其中[N2]-[N3]包含YPAEVVQK(SEQ ID NO:943)、YPPSLVQK(SEQ ID NO:946)、NKAEVVQK(SEQ ID NO:947)、YTAEVVQK(SEQ ID NO:948)、YPAEVEQK(SEQ ID NO:949)、YPAEEVQK(SEQ ID NO:950)、YQAEVVQK(SEQ ID NO:

951)、YTPSLVQK(SEQ ID NO:952)、YPPSLEQK(SEQ ID NO:953)、YPPSLVKK(SEQ ID NO:954)、YPAEVVKK(SEQ ID NO:955)、YPAEVVHK(SEQ ID NO:956)、YPAAVVQK(SEQ ID NO:957)、NPAEVVQK(SEQ ID NO:958)、YPAEVVQQ(SEQ ID NO:959)、CPAEVVQK(SEQ ID NO:960)、YTAEVVKK(SEQ ID NO:961)、YPAEVLQK(SEQ ID NO:962)或YQAEEVQK(SEQ ID NO:963)。

5.如权利要求1至4中任一项所述的AAV衣壳变体,其进一步包含[N1],其中[N1]包含位置X_D、X_E和X_F,其中:

(a)位置X_D为Q、T、S、A、I、L或H;

(b)位置X_E为S、G、A或R;且

(c)位置X_F为S、K、L、R、A或T。

6.如权利要求5所述的AAV衣壳变体,其中[N1]:

(a)包含SK、SL、SS、SR、GA、GS、AS、ST、RS、QS、TS、AG、IG、QA、LG、HS、LS或QR;和/或

(b)包含QSS、QSK、TSL、SSS、QSR、AGA、IGS、QAS、ASS、LGS、QST、HSS、LSS或QRS。

7.如权利要求5或6所述的AAV衣壳变体,其中[N1]-[N2]包含QSSYPAEV(SEQ ID NO:96)、QSKYPAEV(SEQ ID NO:97)、TSLYPAEV(SEQ ID NO:98)、SSSYPAEV(SEQ ID NO:99)、QSRYPAEV(SEQ ID NO:100)、QSSYPPSL(SEQ ID NO:101)、AGAYPAEV(SEQ ID NO:102)、IGSYPAEV(SEQ ID NO:103)、QASYPAEV(SEQ ID NO:104)、ASSYPAEV(SEQ ID NO:105)、LGSYPAEV(SEQ ID NO:106)、QSTNKAEV(SEQ ID NO:107)、HSSYPAEV(SEQ ID NO:108)、SSSYTAEV(SEQ ID NO:109)、TSLYPAEE(SEQ ID NO:110)、ASSYQAEV(SEQ ID NO:111)、QSSYTPSL(SEQ ID NO:112)、QSRYPAEE(SEQ ID NO:113)、LSSYQAEV(SEQ ID NO:114)、HSSYPAAV(SEQ ID NO:115)、QSSNPAEV(SEQ ID NO:116)、QSSYTAEV(SEQ ID NO:117)、TSLCPAEV(SEQ ID NO:118)、QRSYTAEV(SEQ ID NO:119)或QSSYQAEE(SEQ ID NO:120)。

8.如权利要求5至7中任一项所述的AAV衣壳变体,其中:

(a)[N1]-[N2]-[N3]包含SSYPAEVVQ(SEQ ID NO:121)、SKYPAEVVQ(SEQ ID NO:122)、SLYPAEVVQ(SEQ ID NO:123)、SRYPAEVVQ(SEQ ID NO:124)、SSYPPSLVQ(SEQ ID NO:125)、GAYPAEVVQ(SEQ ID NO:126)、GSYPAEVVQ(SEQ ID NO:127)、ASYPAEVVQ(SEQ ID NO:128)、STNKAEVVQ(SEQ ID NO:129)、SSYTAEVVQ(SEQ ID NO:130)、SKYPAEVEQ(SEQ ID NO:131)、SLYPAEEVQ(SEQ ID NO:132)、SSYQAEVVQ(SEQ ID NO:133)、SSYTPSLVQ(SEQ ID NO:134)、SRYPAEEVQ(SEQ ID NO:135)、SSYPPSLEQ(SEQ ID NO:136)、SSYPPSLVK(SEQ ID NO:140)、SSYPAEVVK(SEQ ID NO:141)、SKYPAEVVH(SEQ ID NO:142)、SSYPAAVVQ(SEQ ID NO:143)、SSNPAEVVQ(SEQ ID NO:144)、SLCPAEVVQ(SEQ ID NO:145)、RSYTAEVVQ(SEQ ID NO:146)、SSYTAEVVK(SEQ ID NO:147)、SSYPAEVLQ(SEQ ID NO:148)或SSYQAEEVQ(SEQ ID NO:149);或

(b)[N1]-[N2]-[N3]包含QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、QSSYPPSLVQK(SEQ ID NO:155)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)、QSTNKAEVVQK(SEQ ID NO:161)、HSSYPAEVVQK(SEQ ID NO:162)、SSSYTAEVVQK(SEQ ID NO:163)、QSKYPAEVEQK(SEQ ID NO:164)、TSLYPAEEVQK(SEQ ID NO:165)、ASSYQAEVVQK(SEQ ID NO:166)、QSSYTPSLVQK(SEQ ID NO:167)、QSRYPAEEVQK(SEQ ID

NO:168)、QSSYPPSLEQK(SEQ ID NO:169)、QSSYPPSLVKK(SEQ ID NO:170)、LSSYQAEVVQK(SEQ ID NO:171)、SSSYPAEVVKK(SEQ ID NO:172)、QSKYPAEVVHK(SEQ ID NO:173)、HSSYPAAVVQK(SEQ ID NO:174)、QSSNPAEVVQK(SEQ ID NO:175)、SSSYPAEVVQQ(SEQ ID NO:176)、QSSYTAEVVQK(SEQ ID NO:177)、TSLCPAEVVQK(SEQ ID NO:178)、QRSYTAEVVQK(SEQ ID NO:179)、QSSYTAEVVKK(SEQ ID NO:180)、HSSYPAEVLQK(SEQ ID NO:181)或QSSYQAEEVQK(SEQ ID NO:182)。

9.如权利要求1至8中任一项所述的AAV衣壳变体,其进一步包含[NO],其中[NO]包含位置X_A、X_B和X_C,其中:

(a)位置X_A为T、I或N;

(b)位置X_B为N;且

(c)位置X_C为N、T、S或K。

10.如权利要求9所述的AAV衣壳变体,其中[NO]:

(a)包含TN、IN、NN、NT、NS或NK;或

(b)包含TNN、TNT、INN、TNS、NNN或TNK。

11.如权利要求9或10所述的AAV衣壳变体,其中[NO]-[N1]包含TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNQST(SEQ ID NO:193)、TNNHSS(SEQ ID NO:194)、TNNQSK(SEQ ID NO:184)、TNNLSS(SEQ ID NO:195)、INNQSS(SEQ ID NO:196)、TNSQSS(SEQ ID NO:197)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)、TNNQRS(SEQ ID NO:200)或TNKQAS(SEQ ID NO:201)。

12.如权利要求9至11中任一项所述的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVVQK(SEQ ID NO:510)、TNNQSSYPPSLVQK(SEQ ID NO:512)、TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQK(SEQ ID NO:517)、TNTASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNQSTNKAEVVQK(SEQ ID NO:524)、TNNHSSYPAEVVQK(SEQ ID NO:525)、TNNSSSYTAEVVQK(SEQ ID NO:526)、TNNQSKYPAEVEQK(SEQ ID NO:529)、TNNTSLYPAEEVQK(SEQ ID NO:530)、TNTASSYQAEVVQK(SEQ ID NO:531)、TNNQSSYTPSLVQK(SEQ ID NO:533)、TNNQSRYPAEEVQK(SEQ ID NO:534)、TNNQSSYPPSLEQK(SEQ ID NO:535)、TNNQSSYPPSLVKK(SEQ ID NO:536)、TNNLSSYQAEVVQK(SEQ ID NO:539)、TNNSSSYPAEVVKK(SEQ ID NO:540)、TNNQSKYPAEVVHK(SEQ ID NO:542)、INNQSSYPAEVVQK(SEQ ID NO:543)、TNNHSSYPAAVVQK(SEQ ID NO:545)、TNSQSSNPAEVVQK(SEQ ID NO:548)、TNNSSSYPAEVVQQ(SEQ ID NO:551)、NNNQSRYPAEVVQK(SEQ ID NO:552)、TNNQSSYTAEVVQK(SEQ ID NO:553)、TNNTSLCPAEVVQK(SEQ ID NO:554)、TNSTSLYPAEVVQK(SEQ ID NO:556)、TNNQRSYTAEVVQK(SEQ ID NO:557)、TNNQSSYTAEVVKK(SEQ ID NO:558)、TNNHSSYPAEVLQK(SEQ ID NO:560)、TNNQSSYQAEEVQK(SEQ ID NO:562)或TNKQASYPAEVVQK(SEQ ID NO:563)。

13.如权利要求1至12中任一项所述的AAV衣壳变体,其进一步包含[N4],其中[N4]包含位置X_G和X_H,其中:

(a)位置X_G为T、P或N;且

(b)位置X_H为A。

14.如权利要求13所述的AAV衣壳变体,其中[N4]包含TA、PA或NA。

15.如权利要求13或14所述的AAV衣壳变体,其中[N3]-[N4]包含VQKTA(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKKTA(SEQ ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SEQ ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571)。

16.如权利要求13至15中任一项所述的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]-[N4]包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TNNTSLYP AEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEVVQKTA(SEQ ID NO:1539)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)、TN NQSSYPPSLVQKTA(SEQ ID NO:1300)、TNNAGAYPAEVVQKT A(SEQ ID NO:1021)、TNNIGSYPAEVVQKTA(SEQ ID NO:1112)、TNNQASYPAEVVQKTA(SEQ ID NO:1194)、TNTASSYPAE VVQKTA(SEQ ID NO:1575)、TNNLGSYPAEVVQKTA(SEQ ID NO:1027)、TNNQSTNKAEVVQKTA(SEQ ID NO:1578)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TNNSSSYTAEVVQKTA(SEQ ID NO:1214)、TNNQSKYPAEVEQKTA(SEQ ID NO:1254)、TNNTSLYPAEEVQKTA(SEQ ID NO:1583)、TNTASSYQAEVVQKTA(SEQ ID NO:1584)、TNNQSSYTPSLVQKTA(SEQ ID NO:1585)、TNNQS RYPAEEVQKTA(SEQ ID NO:1342)、TNNQSSYPPSLEQKTA(SEQ ID NO:1590)、TNNQSSYPPSLVKKTA(SEQ ID NO:1591)、TNNLSSYQAEVVQKTA(SEQ ID NO:1592)、TNNQSSYPPSLVQ KPA(SEQ ID NO:1593)、TNNSSSYPAEVVKKTA(SEQ ID NO:1331)、TNNQSKYPAEVVHKTA(SEQ ID NO:1453)、TNNSSSYP AEVVQKPA(SEQ ID NO:1142)、INNQSSYPAEVVQKTA(SEQ ID NO:1024)、TNNHSSYPAAVVQKTA(SEQ ID NO:1598)、TNSQSSNPAEVVQKTA(SEQ ID NO:1599)、TNNSSSYPAEVVQQTA(SEQ ID NO:1419)、NNNQSRYPAEVVQKTA(SEQ ID NO:1601)、TNNQSSYTAEVVQKNA(SEQ ID NO:1602)、TNNTSLCPAEVVQKTA(SEQ ID NO:1603)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)、TNNQRSYTAEVVQKTA(SEQ ID NO:1604)、TNNQ SSYTAEVVKKTA(SEQ ID NO:1606)、TNNHSSYPAEVLQKTA(SEQ ID NO:1607)、TNNQSSYQAEEVQKTA(SEQ ID NO:1608)或TNKQASYPAEVVQKTA(SEQ ID NO:1587)。

17.如权利要求1至16中任一项所述的AAV衣壳变体,其中:

(a)[N2]-[N3]存在于环VIII中;和/或

(b)[NO]、[N1]和[N4]存在于环VIII中;

其中环VIII包含SEQ ID NO:982的位置571-599。

18.如权利要求1至17中任一项所述的AAV衣壳变体,其中:

(i)[NO]的X_A存在于位置571处,[NO]的X_B存在于位置572处,且[NO]的X_C存在于位置573处,这些位置根据SEQ ID NO:982编号;

(ii)[N1]的X_D存在于位置574处,[N1]的X_E存在于位置575处,且[N1]的X_F存在于位置576处,这些位置根据SEQ ID NO:982编号;

(iii)[N2]的X1存在于位置577处,[N2]的X2存在于位置578处,[N2]的X3存在于位置579处,[N2]的X4存在于位置580处,且[N2]的X5存在于位置581处,这些位置根据SEQ ID NO:982编号;

(iv)[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号;和/或

(v)[N4]的X_G存在于位置585处且[N4]的X_H存在于位置586处,这些位置根据SEQ ID NO:982编号。

19.一种AAV衣壳变体,其包含:

(a)表1、2A、2B、9或15-20中所提供的任一序列的氨基酸序列;

(b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或

(c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

20.如权利要求1至19中任一项所述的AAV衣壳变体,其包含:

(a)SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列;

(b)包含来自SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或

(c)相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

21.如权利要求19或20所述的AAV衣壳变体,其中:

(i)3个连续氨基酸包含YPA;

(ii)4个连续氨基酸包含YPAE(SEQ ID NO:21);

(iii)5个连续氨基酸包含YPAEV(SEQ ID NO:1);

(iv)6个连续氨基酸包含YPAEVV(SEQ ID NO:725);

(v)7个连续氨基酸包含YPAEVVQ(SEQ ID NO:726);和/或

(vi)所述氨基酸序列包含YPAEVVQK(SEQ ID NO:943)。

22.如权利要求19至21中任一项所述的AAV衣壳变体,其包含氨基酸序列YPAEVVQK(SEQ ID NO:943)或相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

23.如权利要求1至22中任一项所述的AAV衣壳变体,其中:

(i)所述AAV衣壳变体包含由以下编码的氨基酸序列:SEQ ID NO:944的核苷酸序列;或相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;和/或

(ii)编码所述衣壳变体的核苷酸序列包含SEQ ID NO:944的核苷酸序列;或相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。

24.如权利要求1至23中任一项所述的AAV衣壳变体,其中氨基酸Y存在于位置577处,且PAEVVQK(SEQ ID NO:20)的氨基酸序列存在于位置578-584处,这些位置相对于SEQ ID NO:

982编号。

25.如前述权利要求中任一项所述的AAV衣壳变体,其包含SEQ ID NO:739的氨基酸序列。

26.如前述权利要求中任一项所述的AAV衣壳变体,所述AAV衣壳变体包含SEQ ID NO:738的氨基酸序列,或与其具有至少95%序列同一性的氨基酸序列,任选地,其中所述AAV衣壳变体包含SEQ ID NO:738的氨基酸序列。

27.如权利要求1至26中任一项所述的AAV衣壳变体,其中:

(i)所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少95%序列同一性的氨基酸序列,任选地,其中所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列;和/或

(ii)编码所述AAV衣壳变体的核苷酸序列包含SEQ ID NO:984的核苷酸序列,或与其具有至少90%序列同一性的核苷酸序列,任选地,其中所述核苷酸序列为密码子优化的。

28.一种AAV衣壳变体,其包含SEQ ID NO:982的氨基酸序列。

29.如权利要求1至28中任一项所述的AAV衣壳变体,其中所述AAV衣壳变体

(i)相对于包含SEQ ID NO:138、139或982的氨基酸序列的参考序列的趋向性,对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性;

(ii)转导脑区,例如颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团和/或小脑,任选地,其中如与SEQ ID NO:139的参考序列相比,转导水平大至少1.5、2.2、2.4、2.5、2.6、2.7、3.0、3.2、3.5、3.7、4.0、4.2、4.5、4.7、4.9、5、10、15、20、25、30、35倍,例如当通过例如免疫组织化学测定或qPCR测定的测定测量时,例如,如实施例2中所述;

(iii)与SEQ ID N O:138的参考序列相比,在脑中富集至少约2、3、4、5、6、7、8、9、10、12、15、17、20、25、30、35、40、45、50、55、60或65倍,例如当通过如实施例1中所述的测定测量时;

(iv)与SEQ ID NO:138的参考序列相比,在脑中富集至少约10、12、15、17、20、25、30、35、40、45、50、55、60、61、62、63、64或65倍,例如当通过如实施例1中所述的测定测量时;

(v)与SEQ ID NO:138的参考序列相比,在至少两个至三个物种(例如,非人灵长类动物和啮齿动物(例如,大鼠或小鼠))的脑中富集至少约2、3、4、5、6、7、8、9、10、12、15、17、20、25、30、35、40、45倍,例如当通过如实施例1、2和4中所述的测定测量时;

(vi)与SEQ ID NO:982的参考序列相比,在脑中富集至少约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、100、125、150、175、200或225倍,例如当通过如实施例4中所述的测定测量时;

(vii)将提高水平的有效载荷递送至脑区,任选地,其中如与SEQ ID NO:139的参考序列相比,所述有效载荷的水平提高至少20、25、30、35倍,例如当通过例如qRT-PCR或qPCR测定的测定测量时(例如,如实施例2中所述),任选地,其中所述脑区为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团和/或小脑;

(viii)将提高水平的病毒基因组递送至脑区,任选地,其中如与SEQ ID NO:139的参考序列相比,病毒基因组的水平提高至少1.5、2.2、2.4、2.5、2.6、2.7、3.0、3.2、3.5、3.7、4.0、4.2、4.5、4.7、4.9或5倍,例如当通过例如qRT-PCR或qPCR测定的测定测量时(例如,如实施

例2中所述),任选地,其中所述脑区为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团和/或小脑;

(ix)与SEQ ID NO:139的参考序列相比,在脊髓区中富集至少约3、3.5、4.0、4.5、5、5.0、6.0或6.5倍,例如当通过如实施例2中所述的测定测量时,任选地,其中所述脊髓区为颈区、腰区、胸区或其组合;

(x)相对于背根神经节(DRG)中的转导,显示出脑区中的优先转导;

(xi)相对于肝中的转导,显示出脑区中的优先转导;和/或

(xii)能够转导神经元细胞和非神经元细胞,例如神经胶质细胞(例如,少突胶质细胞或星形胶质细胞)。

30.一种多核苷酸,其编码如权利要求1至29中任一项所述的AAV衣壳变体。

31.如权利要求30所述的多核苷酸,其包含:

(i)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;

(iii)SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;和/或

(iv)SEQ ID NO:984的核苷酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的核苷酸序列,

任选地,其中所述多核苷酸包含密码子优化的核苷酸序列。

32.一种肽,其包含:

(a)表1、2A、2B、9或15-20中所提供的任何序列的氨基酸序列;

(b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或

(c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

33.一种肽,其包含:

(i)YPAEVVQK(SEQ ID NO:943)的氨基酸序列;

(ii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

(iii)来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5、6或7个连续氨基酸。

34.一种AAV颗粒,其包含如权利要求1至29中任一项所述的AAV衣壳变体、由如权利要求30或31所述的多核苷酸编码的AAV衣壳变体或包含如权利要求32或33所述的肽的AAV衣壳变体。

35.如权利要求34所述的AAV颗粒,其包含编码有效载荷的核苷酸序列,任选地,其中所编码的有效载荷包含治疗性蛋白质或其功能变体;抗体或抗体片段;酶;基因编辑系统的组分;RNAi剂(例如,dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA);或其组合。

36.如权利要求35所述的AAV颗粒,其中:

(i)所述治疗性蛋白质或其功能变体(例如,重组蛋白)与神经或神经退行性病症、肌肉或神经肌肉病症或神经肿瘤性病症相关(例如,在其中异常表达),任选地,其中所述治疗性

蛋白质或其功能变体选自载脂蛋白E(APOE)(例如,ApoE2、ApoE3和/或ApoE4);人类运动神经元存活蛋白(SMN)1或SMN2;葡糖脑苷脂酶(GBA1);芳族L-氨基酸脱羧酶(AADC);天冬氨酸酰化酶(ASPA);三肽基肽酶I(CLN2);β-半乳糖苷酶(GLB1);N-磺基葡糖胺磺基水解酶(SGSH);N-乙酰基-α-氨基葡糖苷酶(NAGLU);艾杜糖2-硫酸酯酶(IDS);细胞内胆固醇转运蛋白(NPC1);巨轴索神经蛋白(GAN);或其组合;

(ii)所述抗体或抗体结合片段结合至

(a)CNS相关靶标,例如与神经或神经退行性病症相关的抗原,例如β-淀粉样蛋白、APOE、tau、SOD1、TDP-43、亨廷顿蛋白(HTT)和/或突触核蛋白;

(b)肌肉或神经肌肉相关靶标,例如与肌肉或神经肌肉病症相关的抗原;或

(c)神经肿瘤学相关靶标,例如与神经肿瘤性病症相关的抗原,例如HER2或EGFR(例如,EGFRvIII);

(iii)所述酶包含大范围核酸酶、锌指核酸酶、TALEN、重组酶、整合酶、碱基编辑器、Cas9或其片段;

(iv)基因编辑系统的所述组分包含CRISPR-Cas系统的一种或多种组分,任选地,其中所述CRISPR-Cas系统的所述一种或多种组分包含Cas9,例如Cas9直系同源物或Cpf1,和单一指导RNA(sgRNA),其中:

(a)所述sgRNA位于所述cas9酶的上游(5');和/或

(b)所述sgRNA位于所述cas9酶的下游(3');和/或

(v)所述RNAi剂(例如,dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA)调节(例如,抑制)CNS相关基因、mRNA和/或蛋白质的表达,任选地,其中所述CNS相关基因选自SOD1、MAPT、APOE、HTT、C9ORF72、TDP-43、APP、BACE、SNCA、ATXN1、ATXN3、ATXN7、SCN1A-SCN5A、SCN8A-SCN11A或其组合。

37.如权利要求34至36中任一项所述的AAV颗粒,其包含病毒基因组,所述病毒基因组包含可操作地连接至编码所述有效载荷的核酸序列的启动子,任选地其中:

(i)所述启动子选自人类延伸因子1α-亚单位(EF1α)、巨细胞病毒(CMV)即刻早期增强子和/或启动子、鸡β-肌动蛋白(CBA)及其衍生物CAG、β葡糖醛酸酶(GUSB)或泛素C(UBC)、神经元特异性烯醇化酶(NSE)、血小板源性生长因子(PDGF)、血小板源性生长因子B链(PDGF-β)、细胞间粘附分子2(ICAM-2)、突触蛋白(Syn)、甲基-CpG结合蛋白2(MeCP2)、Ca2+/钙调蛋白依赖性蛋白激酶II(CaMKII)、代谢型谷氨酸受体2(mGluR2)、神经丝轻链(NFL)或重链(NFH)、β-球蛋白微基因nβ2、前脑啡肽原(PPE)、脑啡肽(Enk)和兴奋性氨基酸转运蛋白2(EAAT2)、神经胶质纤维酸性蛋白(GFAP)、髓磷脂碱性蛋白(MBP)、心血管启动子(例如,αMHC、cTnT及CMV-MLC2k)、肝启动子(例如,hAAT、TBG)、骨骼肌启动子(例如,结蛋白、MCK、C512)或片段(例如,截短)或其功能变体;

(ii)所述启动子为EF-1a启动子变体,例如截短的EF-1a启动子;或

(iii)所述启动子包含SEQ ID NO:987-1007中任一个的核苷酸序列;相对于SEQ ID NO:987-1007的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过四个修饰(例如,取代)的核苷酸序列;或与SEQ ID NO:987-1007中任一个具有至少80%(例如,85%、90%、95%、96%、97%、98%或99%)序列同一性的核苷酸序列。

38.如权利要求37所述的AAV颗粒,其中所述病毒基因组进一步包含:

(i)polyA信号序列;

(ii)反向末端重复(ITR)序列,任选地,其中所述ITR序列相对于所编码的有效载荷位于5'处和/或所述ITR序列相对于所编码的有效载荷位于3'处;

(iii)增强子、Kozak序列、内含子区和/或外显子区;和/或

(iv)编码miR结合位点的核苷酸序列,所述miR结合位点例如调节(例如,减少)由所述病毒基因组编码的抗体分子在其中表达相应miRNA的细胞或组织中的表达的miR结合位点,任选地,其中所编码的miR结合位点调节(例如,减少)所编码的抗体分子在DRG、肝、心脏、造血谱系或其组合的细胞或组织中的表达。

39.如权利要求38所述的AAV颗粒,其中所述病毒基因组包含:

(i)所编码的miR结合位点的至少1-5个拷贝,例如至少1、2、3、4或5个拷贝;

(ii)所编码的miR结合位点的至少3个拷贝,任选地其中:

(a)所有三个拷贝包含相同miR结合位点,或这些拷贝中的至少一个、两个、三个或全部包含不同miR结合位点;和/或

(b)所编码的miR结合位点的所述3个拷贝为连续的(例如,

未由间隔子分开)或由间隔子分开,任选地,其中所述间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,具有至少一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的核苷酸序列;或

(iii)所编码的miR结合位点的至少4个拷贝,任选地其中

(a)所有四个拷贝均包含相同miR结合位点,或这些拷贝中的至少一个、两个、三个或全部包含不同miR结合位点;和/或

(b)所编码的miR结合位点的所述4个拷贝为连续的(例如,

未由间隔子分开)或由间隔子分开,任选地,其中所述间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,具有至少一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的核苷酸序列。

40.如权利要求38或39所述的AAV颗粒,其中所编码的miR结合位点包含miR122结合位点、miR183结合位点、miR-1结合位点、miR-142-3p或其组合,任选地其中:

(i)所编码的miR122结合位点包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

(ii)所编码的miR183结合位点包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

(iii)所编码的miR-1结合位点包含SEQ ID NO:4679的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4679,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;和/或

(iv)所编码的miR-142-3p结合位点包含SEQ ID NO:4675的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4675,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

41.如权利要求37至40中任一项所述的AAV颗粒,其中所述病毒基因组为:

(i)单链的;或

(ii)自身互补的。

42.如前述权利要求中任一项所述的AAV衣壳变体、多核苷酸、肽或AAV颗粒,其为分离的,例如重组的。

43.一种载体,其包含编码如权利要求1至29或42中任一项所述的AAV衣壳变体的多核苷酸、如权利要求30、31或42中任一项所述的多核苷酸或编码如权利要求32、33或42中任一项所述的肽的多核苷酸。

44.一种细胞,例如宿主细胞,所述细胞包含如权利要求1至29或42中任一项所述的AAV衣壳变体、如权利要求30、31或42中任一项所述的多核苷酸、编码如权利要求32、33或42中任一项所述的肽的多核苷酸、如权利要求34至42中任一项所述的AAV颗粒或如权利要求43所述的载体,任选地其中:

(i)所述细胞为哺乳动物细胞或昆虫细胞;

(ii)所述细胞为脑区或脊髓区的细胞,任选地为脑干、海马体或丘脑的细胞;和/或

(iii)所述细胞为神经元、感觉神经元、运动神经元、星形胶质细胞、神经胶质细胞、少突胶质细胞或肌肉细胞(例如,心脏、横膈膜或股四头肌的细胞)。

45.一种制备AAV颗粒的方法,其包括

(i)提供包含病毒基因组的宿主细胞;及

(ii)在适合将所述病毒基因组封入如权利要求1至29或42中任一项所述的AAV衣壳变体或由如权利要求30、31或32中任一项所述的多核苷酸编码的AAV衣壳变体中的条件下培育所述宿主细胞;

由此制备所述AAV颗粒。

46.一种药物组合物,其包含如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒,以及药学上可接受的赋形剂。

47.一种将有效载荷递送至细胞或组织(例如,CNS细胞或CNS组织)的方法,其包括施用有效量的如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒。

48.如权利要求47所述的方法,其中所述细胞为:

(i)脑区或脊髓区的细胞,任选地为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑、脊髓颈区、脊髓胸区、脊髓腰区或其组合的细胞;

(ii)为心脏(例如,心房或心室)的细胞;

(iii)为肌肉细胞(例如,股四头肌的细胞)或肝细胞;

(iv)神经元、感觉神经元、运动神经元、星形胶质细胞、神经胶质细胞或少突胶质细胞;

(v)在受试者内,任选地,其中所述受试者患有、已被诊断患有遗传病症(例如,单基因病症或多基因病症)、神经病症(例如,神经退行性病症)、神经肿瘤性病症、肌肉病症或神经肌肉病症,或处于患有所述病症的风险中。

49.一种治疗患有或被诊断患有遗传病症(例如,单基因病症或多基因病症)的受试者的方法,所述方法包括向所述受试者施用有效量的如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒。

50.一种治疗患有或被诊断患有神经病症(例如,神经退行性病症)、神经肿瘤性病症、肌肉病症或神经肌肉病症的受试者的方法,所述方法包括向所述受试者施用有效量的如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒。

51.一种治疗患有或被诊断患有肌肉病症或神经肌肉病症的受试者的方法,所述方法包括向所述受试者施用有效量的如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒、包含如权利要求32、33或42中任一项所述的肽的AAV颗粒。

52.一种治疗患有或被诊断患有神经肿瘤性病症的受试者的方法,所述方法包括向所述受试者施用有效量的如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒。

53.如权利要求48至52中任一项所述的方法,其中所述遗传病症、神经病症、神经退行性病症、肌肉病症、神经肌肉病症或神经肿瘤性病症为亨廷顿氏病、肌萎缩性侧索硬化(ALS)、戈谢病、路易体痴呆、帕金森氏病、脊髓性肌萎缩、阿尔茨海默氏病、脑白质营养不良(例如,亚历山大病、常染色体显性脑白质营养不良伴自主神经疾病(ADLD)、卡纳万病、脑腱黄瘤病(CTX)、异染性脑白质营养不良(MLD)、佩-梅二氏病或雷夫叙姆病)或癌症(例如,HER2/neu阳性癌症或胶质母细胞瘤)。

54.如权利要求49至53中任一项所述的方法,其中治疗包括预防所述受试者的所述疾病或病症的进展。

55.如权利要求48至54中任一项所述的方法,其中所述受试者为人。

56.如权利要求48至55中任一项所述的方法,其中经由以下途径向所述受试者施用所述AAV颗粒:

(i)静脉内、经由大池内注射(ICM)、脑内、鞘内、脑室内、经由实质内施用或肌肉内;

(ii)经由聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合;或

(iii)静脉内。

57.如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33

或42中任一项所述的肽的AAV颗粒,其用于将有效载荷递送至细胞或组织的方法中。

58.如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒,其用于治疗遗传病症、神经病症、神经退行性病症、肌肉病症、神经肌肉病症或神经肿瘤性病症的方法中。

59.如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒,其用于药剂的制造。

60.如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒在药剂的制造中的用途。

61.如权利要求46所述的药物组合物、如权利要求34至42中任一项所述的AAV颗粒、包含如权利要求1至29或42中任一项所述的AAV衣壳变体的AAV颗粒或包含如权利要求32、33或42中任一项所述的肽的AAV颗粒在用于治疗遗传病症、神经病症、神经退行性病症、肌肉病症、神经肌肉病症或神经肿瘤性病症的药剂的制造中的用途。

AAV衣壳变体及其用途

[0001] 相关申请

[0002] 本申请要求2022年2月8日提交的美国临时申请号63/307,742和2022年5月9日提交的美国临时申请号63/339,574的优先权,其各自的完整内容以引用方式整体并入本文中。

[0003] 序列表

[0004] 本申请含有序列表,所述序列表已以XML格式以电子方式提交且以引用方式整体并入本文中。所述XML拷贝创建于2023年1月24日,命名为V2071-1111PCT_SL.xml且大小为2,233,613个字节。

技术领域

[0005] 本公开涉及用于制备、使用和/或配制腺相关病毒衣壳蛋白及其变体的组合物和方法。

背景技术

[0006] 基因递送至中枢神经系统(CNS)仍是基因疗法中的重大挑战。具有改善的脑趋向性的工程改造的腺相关病毒(AAV)衣壳代表一种有吸引力的解决CNS递送限制的方法。

[0007] AAV源性载体由于其非致病性质、其低免疫原性型态、整合到宿主基因组中的低速率以及在非分裂细胞中的长期转基因表达而为用于临床基因转移的有希望的工具。然而,AAV天然变体在某些器官中的转导效率对于临床应用来说过低,且通过预存在的中和抗体进行的衣壳中和可阻止大部分患者的治疗。由于这些原因,已付出相当大的努力来获得具有增强特性的衣壳变体。在迄今为止所测试的许多方法中,重大进展来自使用衣壳变体的体外或体内选择进行的AAV衣壳的定向演化,这些变体通过使用易错PCR、各种亲本血清型的改组或在规定位置处插入完全随机化的短肽进行的衣壳序列随机化而产生。

[0008] 已尝试提供具有改善特性的AAV衣壳,例如,在全身性施用后对靶细胞或组织的趋向性被改进,但取得的成效有限。因而,需要改进的方法来产生AAV衣壳且将所得AAV衣壳的相关有效载荷递送至靶细胞或组织,例如CNS细胞或组织或者肌肉细胞或组织。

发明内容

[0009] 本公开至少部分涉及用于产生和使用包含AAV衣壳多肽(例如,AAV衣壳变体(例如,AAV5衣壳变体))的AAV颗粒的组合物和方法。在一些实施方案中,AAV衣壳变体对组织或细胞(例如,CNS组织、CNS细胞、心脏细胞、心脏组织、肌肉细胞、肌肉组织、肝细胞或肝组织)具有增强的趋向性。该趋向性可用于将有效载荷(例如,本文所述的有效载荷)递送至细胞或组织,以治疗病症,例如神经或神经退行性病症、肌肉或神经肌肉病症或神经肿瘤性病症。

[0010] 因此,一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),其包含相对于SEQ ID NO:138编号的位置577处非T的氨基酸(例如,Y、N或C)。在一些实施方案中,AAV衣壳

变体包含相对于SEQ ID NO:138编号的位置577处的Y。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:138编号的位置577处的N。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:138编号的位置577处的C。

[0011] 另一方面,本公开提供一种AAV衣壳变体(例如,野生型AAV5衣壳的变体),所述AAV衣壳变体包含超过一个置换相对于SEQ ID NO:138编号的位置577处的苏氨酸(T)的氨基酸。在一些实施方案中,两个、三个、四个、五个、六个、七个、八个、九个或十个氨基酸的插入物置换相对于SEQ ID NO:138编号的位置577处的T。在一些实施方案中,八个氨基酸的插入物置换相对于SEQ ID NO:138编号的位置577处的T。

[0012] 另一方面,本公开提供一种包含式[N2]-[N3]的AAV衣壳变体(例如,AAV5衣壳变体),其中:(i)[N2]包含位置X1、X2、X3、X4和X5,其中:(a)位置X1为Y、N或C;(b)位置X2为P、K、T或Q;(c)位置X3为A或P;(d)位置X4为E、S或A;且(e)位置X5为V、L或E;和(ii)[N3]的氨基酸序列包含VQK、EQK、VKK、VHK、VQQ或LQK的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中的任一上述氨基酸的氨基酸修饰(例如,保守取代);任选地,其中[N2]-[N3]紧接在相对于SEQ ID NO:138编号的位置576之后存在且置换相对于SEQ ID NO:138编号的位置577。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N3]的氨基酸序列为VQK。在一些实施方案中,[N2]的X1存在于位置577处,[N2]的X2存在于位置578处,[N2]的X3存在于位置579处,[N2]的X4存在于位置580处,且[N2]的X5存在于位置581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。

[0013] 在一些实施方案中,[N2]的氨基酸序列由YPAEV(SEQ ID NO:1)组成。在一些实施方案中,[N3]的氨基酸序列由VQK组成。在一些实施方案中,[N2]-[N3]置换野生型AAV5(例如,SEQ ID NO:138)的位置577处的苏氨酸(T)。在一些实施方案中,[N2]-[N3]置换野生型AAV5(例如,SEQ ID NO:138)的位置577处的苏氨酸(T),且[N2]的氨基酸序列由YPAEV(SEQ ID NO:1)组成,且[N3]的氨基酸序列由VQK组成。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。

[0014] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个或全部:(i)[N0],其中[N0]的氨基酸序列包含TNN、TNT、INN、TNS、NNN或TNK;(ii)[N1],其中[N1]的氨基酸序列包含QSS、QSK、TSL、SSS、QSR、AGA、IGS、QAS、ASS、LGS、QST、HSS、LSS或QRS;(iii)[N2],其中[N2]的氨基酸序列包含YPAEV(SEQ ID NO:1)、YPPSL(SEQ ID NO:2)、NKAEV(SEQ ID NO:3)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、YQAEV(SEQ ID NO:6)、YTPSL(SEQ ID NO:7)、YPAAV(SEQ ID NO:8)、NPAEV(SEQ ID NO:9)、CPAEV(SEQ ID NO:10)或YQAEE(SEQ ID NO:11);(iv)[N3],其中[N3]的氨基酸序列包含VQK、EQK、VKK、VHK、VQQ或LQK;和/或(v)[N4],其中[N4]的氨基酸序列包含TA、PA或NA;且任

选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代);任选地,其中[N0]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在一些实施方案中,[N0]的氨基酸序列为TNN。在一些实施方案中,[N1]的氨基酸序列为QSS。在一些实施方案中,[N2]的氨基酸序列为YPAEV(SEQ ID NO:1)。在一些实施方案中,[N3]的氨基酸序列为VQK。在一些实施方案中,[N4]的氨基酸序列为TA。在一些实施方案中,[N0]的氨基酸序列为TNN,[N1]的氨基酸序列为QSS,[N2]的氨基酸序列为YPAEV(SEQ ID NO:1),[N3]的氨基酸序列为VQK,和/或[N4]的氨基酸序列为TA。在一些实施方案中,[N0]的氨基酸序列为TNN,[N1]的氨基酸序列为QSS,[N2]的氨基酸序列为YPAEV(SEQ ID NO:1),[N3]的氨基酸序列为VQK,及[N4]的氨基酸序列为TA。在这些实施方案的任一个中,[N0]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,[N0]置换位置571-573(例如,T571、N572及N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N0]紧接在位置570之后存在,且[N0]位置571-573(例如,T571、N572及N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N1]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在这些实施方案的任一个中,[N1]置换位置574-576(例如,Q574、S575及S576),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N1]紧接在位置573之后存在,且[N1]置换位置574-576(例如,Q574、S575及S576),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在这些实施方案的任一个中,[N2]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在这些实施方案的任一个中,[N2]紧接在位置576之后存在,且[N2]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]-[N3]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在这些实施方案的任一个中,[N2]-[N3]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在这些实施方案的任一个中,[N2]-[N3]紧接在位置576之后存在,且[N2]-[N3]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578及A579),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]-[N3]-[N4]紧接在位置576之后存在,且[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578及A579),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N0]-[N1]-[N2]-[N3]-[N4]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578及A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]紧接在位置570之后存在,且[N0]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578及A579),这些位置相对于SEQ ID NO:138编号。例如,在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)且紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中[N2]-[N3](YPAEVVQK(SEQ ID NO:943))置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。在一些实施方案中,[N0]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N1]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982

编号。在一些实施方案中,[N4]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0015] 另一方面,本公开提供一种包含式[B]-[C]的AAV衣壳变体(例如,AAV5衣壳变体),其中:(i)[B]包含位置X1、X2和X3,其中:(a)位置X1为Q、T、S、A、I、L或H;(b)位置X2为S、G或A;且(c)位置X3为S、K、L、R或A;和(ii)[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中的任一上述氨基酸的氨基酸修饰(例如,保守取代);且进一步任选地,其中[B]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在一些实施方案中,[B]包含位置X1处的Q。在一些实施方案中,[B]包含位置X2处的S。在一些实施方案中,[B]包含位置X3处的S。在一些实施方案中,[B]的氨基酸序列为QSS。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]置换位置574-576(例如,Q574、S575和S576)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在,且[B]置换位置574-576(例如,Q574、S575和S576)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在,且[C]置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在,且[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在一些实施方案中,[B]的X1存在于位置574处,[B]的X2存在于位置575处,且[B]的X3存在于位置576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]-[C]存在于位置574-584处,这些位置根据SEQ ID NO:982编号。

[0016] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个或全部:(i)[A],其中[A]的氨基酸序列包含TNN、TNT、INN、NNN、TNS或TNK;(ii)[B],其中[B]的氨基酸序列包含QSS、TSL、SSS、QSR、QSK、AGA、IGS、QAS、ASS、LGS或HSS;(iii)[C],其中[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;和(iv)[D],其中[D]包含TA或PA的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代);任选地,其中[A]紧接在相对于SEQ ID NO:138编号的位置570之后存在且[C]置换相对于SEQ ID NO:138编号的位置577。在一些实施方案中,[A]的氨基酸序列为TNN。在一些实施方案中,[B]的氨基酸序列为QSS。在一些实施方案中,[A]的氨基酸序列为TNN且[B]的氨基酸序列为QSS。在一些实施方案中,[A]的氨基酸序列为TNN,[B]的氨基酸序列为QSS,且[C]的氨基酸序列为YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,[A]的氨基酸序列为TNN,[B]的氨基酸序列为QSS,[C]的氨基酸序列为YPAEVVQK

(SEQ ID NO:943),且[D]的氨基酸序列为TA。在这些实施方案的任一个中,[A]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,[A]置换位置571-573(例如,T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[A]紧接在位置570之后存在,且[A]置换位置571-573(例如,T571、N572及N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在。在这些实施方案的任一个中,[B]置换位置574-576(例如,Q574、S575和S576),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[B]紧接在位置573之后存在,且[B]置换位置574-576(例如,Q574、S575及S576),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[C]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在这些实施方案的任一个中,[C]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在这些实施方案的任一个中,[C]紧接在位置576之后存在,且[C]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[B]-[C]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在这些实施方案的任一个中,[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[B]-[C]紧接在位置573之后存在,且[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[C]-[D]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在这些实施方案的任一个中,[C]-[D]置换位置577-579(例如,T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[A]-[B]-[C]-[D]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,[A]-[B]-[C]-[D]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[A]-[B]-[C]-[D]紧接在位置570之后存在,且[A]-[B]-[C]-[D]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。例如,在一些实施方案中,[A]-[B]-[C]-[D]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)且紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中[C](YPAEVVQK(SEQ ID NO:943))置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。在一些实施方案中,[A]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[D]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[A]-[B]-[C]-[D]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0017] 另一方面,本公开提供一种包含式[N2]-[N3]的AAV衣壳变体(例如,AAV5衣壳变体),其中:(i)[N2]包含位置X1、X2、X3、X4和X5,其中:(a)位置X1为Y或T;(b)位置X2为Q、T、P或E;(c)位置X3为A;(d)位置X4为E或D;且(e)位置X5为V或E;和(ii)[N3]包含VQK或VQN的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中的任一上述氨基酸的氨基酸修饰(例如,保守取代);且进一步任选地,其中[N2]-[N3]紧接在相对于SEQ ID NO:138编号的位置576之后存在且置换相对于SEQ ID NO:138编号的位置577。在一些实施方案中,[N2]的氨基酸序列由YPAEV(SEQ ID NO:1)组成。在一些实施方案中,[N3]的氨基酸序列由

VQK组成。在一些实施方案中,[N2]-[N3]置换野生型AAV5(例如,SEQ ID NO:138)的位置577(例如,T577)。在一些实施方案中,[N2]-[N3]置换野生型AAV5(例如,SEQ ID NO:138)的位置577处的苏氨酸(T),且[N2]的氨基酸序列由YPAEV(SEQ ID NO:1)组成,且[N3]的氨基酸序列由VQK组成。在一些实施方案中,[N2]的X1存在于位置577处,[N2]的X2存在于位置578处,[N2]的X3存在于位置579处,[N2]的X4存在于位置580处,且[N2]的X5存在于位置581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。

[0018] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:(i)[NO],其中[NO]包含TNN、TNS、TNT或TNK的氨基酸序列;(ii)[N1],其中[N1]包含QSS、SLS、SLY、SAT或QTS的氨基酸;(iii)[N2],其中[N2]包含YPAEV(SEQ ID NO:1)、YQAEV(SEQ ID NO:6)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、TEAEV(SEQ ID NO:12)或YPADV(SEQ ID NO:13)的氨基酸序列;(iv)[N3],其中[N3]包含VQK或VQN的氨基酸序列;和/或(v)[N4],其中[N4]包含TA、PA、TD、NA或PA的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代);且进一步任选地,其中[NO]紧接在相对于SEQ ID NO:138编号的位置570之后存在且[N2]-[N3]置换相对于SEQ ID NO:138编号的位置577。在这些实施方案的任一个中,[NO]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,[NO]置换位置571-573(例如,T571、N572及N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[NO]紧接在位置570之后存在,且[NO]置换位置571-573(例如,T571、N572及N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N1]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在这些实施方案的任一个中,[N1]置换位置574-576(例如,Q574、S575及S576),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N1]紧接在位置573之后存在,且[N1]置换位置574-576(例如,Q574、S575和S576),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在这些实施方案的任一个中,[N2]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在这些实施方案的任一个中,[N2]紧接在位置576之后存在,且[N2]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]-[N3]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在这些实施方案的任一个中,[N2]-[N3]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在这些实施方案的任一个中,[N2]-[N3]紧接在位置576之后存在,且[N2]-[N3]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[N2]-[N3]-[N4]紧接在位置576之后存在,且[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[NO]-[N1]-[N2]-[N3]-[N4]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[NO]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、

Q574、S575、S576、T577、T578和A579)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[N0]-[N1]-[N2]-[N3]-[N4]紧接在位置570之后存在,且[N0]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。在一些实施方案中,[N0]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N1]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N4]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0019] 在一些实施方案中,[N0]的氨基酸序列为TNN,[N1]的氨基酸序列为QSS,[N2]的氨基酸序列为YPAEV(SEQ ID NO:1),[N3]的氨基酸序列为VQK,且/或[N4]的氨基酸序列为TA。在一些实施方案中,[N0]的氨基酸序列为TNN,[N1]的氨基酸序列为QSS,[N2]的氨基酸序列为YPAEV(SEQ ID NO:1),[N3]的氨基酸序列为VQK,且[N4]的氨基酸序列为TA。在这些实施方案的任一个中,[N0]-[N1]-[N2]-[N3]-[N4]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,[N2]-[N3]的氨基酸序列置换相对于SEQ ID NO:138编号的位置577。例如,在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)且紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中[N2]-[N3](YPAEVVQK(SEQ ID NO:943))置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。在一些实施方案中,[N0]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N1]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N4]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0020] 另一方面,本公开提供一种包含[B]-[C]的AAV衣壳变体(例如,AAV5衣壳变体),其中:(i)[B]包含位置X1、X2和X3,其中(a)位置X1为Q或S;(b)位置X2为S、L或A;且(c)位置X3为S、Y或T;和(ii)[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中的任一上述氨基酸的氨基酸修饰(例如,保守取代);且进一步任选地,其中[B]紧接在相对于SEQ ID NO:138编号的位置573之后存在,且[C]置换相对于SEQ ID NO:138编号的位置577。在一些实施方案中,[B]为QSS。在一些实施方案中,[C]由YPAEVVQK(SEQ ID NO:943)组成,紧接在相对于SEQ ID NO:138编号的位置576之后存在,且置换相对于SEQ ID NO:138编号的位置577。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]置换位置574-576(例如,Q574、S575和S576)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在,且[B]置换位置574-576(例如,Q574、S575和S576)。在一些实施方案中,相对于根据SEQ ID NO:138编号

的参考序列,[C]紧接在位置576之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在,且[C]置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在,且[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在一些实施方案中,[B]的X1存在于位置574处,[B]的X2存在于位置575处,且[B]的X3存在于位置576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]-[C]存在于位置574-584处,这些位置根据SEQ ID NO:982编号。

[0021] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:(i)[A],其中[A]包含TNN、TNS、TNT或TNK的氨基酸序列;(ii)[B],其中[B]包含QSS、SLY、SAT或SLS的氨基酸序列;(iii)[C],其中[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;和(iv)[D],其中[D]包含TA、TD、NA或PA的氨基酸序列;且任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代);且进一步任选地,其中[A]紧接在相对于SEQ ID NO:138编号的位置570之后存在,且[C]置换相对于SEQ ID NO:138编号的位置577。在这些实施方案的任一个中,[A]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在这些实施方案的任一个中,[A]置换位置571-573(例如,T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,[A]紧接在位置570之后存在,且[A]置换位置571-573(例如,T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]置换位置574-576(例如,Q574、S575和S576)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在,且[B]置换位置574-576(例如,Q574、S575和S576)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[C]置换位置577(例如,T577)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在,且[C]置换位置577(例如,T577)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在,且[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[C]-[D]紧接在位置576之后存在。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[C]-[D]置换位置577-579(例如,T577、T578和A579)。在这些实施方案的任一个中,相对于根据SEQ ID NO:138编号的参考序列,[C]-[D]紧接在位置576之后存

在,且[C]-[D]置换位置577-579(例如,T577、T578和A579)。在一些实施方案中,[A]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[D]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[A]-[B]-[C]-[D]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0022] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含[K1]-[K2],其中:(i)[K1]包含LSY或LYY,和(ii)[K2]包含位置X1、X2、X3及X4,其中(a)位置X1为Q、T或P;(b)位置X2为A;(c)位置X3为E或D;且(d)位置X4为V或E;且任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]紧接在位置574之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]置换位置575-577(例如,S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]紧接在位置574之后存在,且[K1]-[K2]置换位置575-577(例如,S575、S576和T577)。

[0023] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:(i)[K0],其包含TNNS(SEQ ID NO:14);(ii)[K1],其包含LSY或LYY;(iii)[K2],其包含QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)或PADV(SEQ ID NO:19);(iv)[K3],其包含VQK或VQN;和(v)[K4],其包含TA、TD、NA或PA;且任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K0]紧接在位置570之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K0]置换位置571-574(例如,T571、N572、N573和Q574)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K0]紧接在位置570之后存在,且[K0]置换位置571-574(例如,T571、N572、N573和Q574)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]紧接在位置574之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]置换位置575-577(例如,S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]紧接在位置574之后存在,且[K1]置换位置575-577(例如,S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]-[K3]紧接在位置574之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]-[K3]置换位置575-577(例如,S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]-[K3]紧接在位置574之后存在,且[K1]-[K2]-[K3]置换位置575-577(例如,S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]-[K3]-[K4]紧接在位置574之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]-[K3]-[K4]置换位置575-579(例如,S575、S576、T577、T578和A579)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K1]-[K2]-[K3]-[K4]紧接在位置574之后存在,且[K1]-[K2]-[K3]-[K4]置换位置575-579(例如,S575、S576、T577、T578和A579)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[K0]-[K1]-[K2]-

[K3]-[K4]紧接在位置570之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,[K0]-[K1]-[K2]-[K3]-[K4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,[K0]-[K1]-[K2]-[K3]-[K4]紧接在位置570之后存在,且[K0]-[K1]-[K2]-[K3]-[K4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和T579)。

[0024] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含:(a)表1、2A、2B、9或15-20中所提供的任何序列的氨基酸序列;(b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;(c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或(d)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0025] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含:(a)SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列;(b)包含来自SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;(c)相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或(d)相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0026] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含(a)SEQ ID NO:943或2064-2080中任一个的氨基酸序列;(b)包含来自SEQ ID NO:943或2064-2080中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;(c)相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或(d)相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0027] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含SEQ ID NO:943的氨基酸序列。在一些实施方案中,相对于根据SEQ ID NO:138或982编号的参考序列,SEQ ID NO:943的氨基酸序列紧接在位置576之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,SEQ ID NO:943的氨基酸序列置换位置577(例如,置换T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,SEQ ID NO:943的氨基酸序列紧接在位置576之后存在,其中SEQ ID NO:943的氨基酸序列置换位置

577(例如,置换T577)。

[0028] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943),该序列紧接在相对于SEQ ID NO:138或982编号的位置576之后存在。在一些实施方案中,YPAEVVQK(SEQ ID NO:943)的氨基酸序列置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[0029] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943),其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[0030] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943),该序列紧接在相对于SEQ ID NO:138编号的位置576之后存在,且其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[0031] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),其包含位置577处的氨基酸Y,且进一步包含PAEVVQK(SEQ ID NO:20)的氨基酸序列,该序列紧接在相对于SEQ ID NO:138编号的位置577之后存在。

[0032] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),其包含位置577处的氨基酸Y,且包含PAEVVQK(SEQ ID NO:20)的氨基酸序列,该序列紧接在相对于SEQ ID NO:982编号的位置577之后存在。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:739的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:738的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:982的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。

[0033] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),其包含位置577处的氨基酸Y且包含位置578-584处的PAEVVQK(SEQ ID NO:20)的氨基酸序列,这些位置相对于SEQ ID NO:982编号。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:739的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:738的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:982的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。

[0034] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533),该序列紧接在相对于SEQ ID NO:138或982编号的位置570之后存在。在一些实施方案中,YPAEVVQK(SEQ ID NO:943)的氨基酸序列置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[0035] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533),其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[0036] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533),该序列紧接在相对于SEQ ID NO:138编号的位

置570之后存在,其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[0037] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533),该序列存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0038] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列。另一方面,本公开提供一种AAV衣壳变体,所述AAV衣壳变体由SEQ ID NO:982的氨基酸序列组成。

[0039] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含由SEQ ID NO:984的核苷酸序列编码的氨基酸序列。另一方面,本公开提供一种AAV衣壳变体,所述AAV衣壳变体包含由与SEQ ID NO:984具有至少95%同一性的核苷酸序列编码的氨基酸序列。

[0040] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含与SEQ ID NO:982的位置193-731的氨基酸序列至少90%、95%、96%、97%、98%或99%同一的氨基酸序列,其中所述AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:982的位置193-731的氨基酸序列。

[0041] 另一方面,本公开提供一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含与SEQ ID NO:739的氨基酸序列至少90%、95%、96%、97%、98%或99%同一的氨基酸序列,其中所述AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:739的氨基酸序列。

[0042] 另一方面,本公开提供一种肽,其包含:(a)表1、2A、2B、9或15-20中所提供的任何序列的氨基酸序列;(b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或(c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或(d)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0043] 另一方面,本公开提供一种肽,所述肽包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列。另一方面,本公开提供一种肽,所述肽包含(i)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;(ii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的氨基酸序列;或(iii)来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5、6或7个连续氨基酸。

[0044] 另一方面,本公开提供一种肽,所述肽由SEQ ID NO:944的核苷酸序列或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列编码。另一方面,本公开提供一种肽,所述肽由以下编码:(i)相对于SEQ ID NO:944的核苷酸序列,包含一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;或(ii)相对于SEQ ID NO:944的核苷酸序列,包含一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入

或缺失)的核苷酸序列。

[0045] 另一方面,本公开提供一种肽,其中编码所述肽的核苷酸序列包含(i)SEQ ID NO:944的核苷酸序列或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;(ii)相对于SEQ ID NO:944的核苷酸序列,包含一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;或(iii)相对于SEQ ID NO:944的核苷酸序列,包含一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[0046] 另一方面,本公开提供一种编码AAV衣壳变体(例如,AAV5衣壳变体)的多核苷酸,其中所编码的AAV衣壳变体包含:(a)表1、2A、2B、9或15-20中所提供的任何序列的氨基酸序列;(b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或(c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或(d)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0047] 另一方面,本公开提供一种编码AAV衣壳变体(例如,AAV5衣壳变体)的多核苷酸,其中所编码的AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列。另一方面,本公开提供一种编码AAV衣壳变体的多核苷酸,其中所编码的AAV衣壳变体包含(i)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;(ii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的氨基酸序列;或(iii)来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5、6或7个连续氨基酸。

[0048] 另一方面,本公开提供一种AAV颗粒,所述AAV颗粒包含本文所述的AAV衣壳多肽,例如AAV衣壳变体(例如,AAV5衣壳变体)。在一些实施方案中,所述AAV颗粒包含编码有效载荷的核酸序列。在一些实施方案中,所述AAV颗粒进一步包含病毒基因组,所述病毒基因组包含可操作地连接至编码有效载荷的核酸序列的启动子。

[0049] 另一方面,本公开提供一种制备AAV颗粒的方法,所述AAV颗粒包含本文所述的AAV衣壳多肽,例如AAV衣壳变体(例如,AAV5衣壳变体)。在一些实施方案中,所述方法包括提供包含病毒基因组的宿主细胞且在适合将病毒基因组封入AAV衣壳变体(例如,本文所述的AAV衣壳变体)中的条件下培育宿主细胞,由此制备所述AAV颗粒。

[0050] 另一方面,本公开提供一种将有效载荷递送至细胞或组织(例如,CNS细胞、CNS组织、肌肉细胞或肌肉组织)的方法。所述方法包括施用有效量的包含本文所述的AAV衣壳变体(例如,AAV5衣壳变体)的AAV颗粒。

[0051] 另一方面,本公开提供一种治疗患有或被诊断患有遗传病症(例如,单基因病症或多基因病症)的受试者的方法。所述方法包括施用有效量的包含本文所述的AAV衣壳变体(例如,AAV5衣壳变体)的AAV颗粒。

[0052] 另一方面,本公开提供一种治疗患有或被诊断患有神经(例如,神经退行性)病症

的受试者的方法。所述方法包括施用有效量的包含本文所述的AAV衣壳变体(例如,AAV5衣壳变体)的AAV颗粒。

[0053] 另一方面,本公开提供一种治疗患有或被诊断患有肌肉病症或神经肌肉病症的受试者的方法。所述方法包括施用有效量的包含本文所述的AAV衣壳变体(例如,AAV5衣壳变体)的AAV颗粒。

[0054] 另一方面,本公开提供一种治疗患有或被诊断患有心脏病症的受试者的方法,例如本文所述的心脏病症(例如,心肌病(例如,致心律失常性右心室心肌病、扩张型心肌病或肥厚性心肌病)、充血性心脏衰竭、心动过速(例如,儿茶酚胺激导性多形性室性心动过速)、缺血性心脏病和/或心肌梗塞)。所述方法包括施用有效量的包含本文所述的AAV衣壳变体(例如,AAV5衣壳变体)的AAV颗粒。

[0055] 另一方面,本公开提供一种治疗患有或被诊断患有神经肿瘤性病症的受试者的方法。所述方法包括施用有效量的包含本文所述的AAV衣壳变体(例如,AAV5衣壳变体)的AAV颗粒。

[0056] 本领域技术人员将认识到或能够仅使用常规实验来确定本文所述的本发明的特定实施方案的许多等效物。此类等效物意图由以下列举的实施方案涵盖。

[0057] 列举的实施方案

[0058] 1.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包括包含下式的胺基序列:

[0059] [N2]-[N3],其中:

[0060] (i)[N2]包含位置X1、X2、X3、X4和X5,其中:

[0061] (a)位置X1为Y、N、C或T;

[0062] (b)位置X2为P、E、K、T或Q;

[0063] (c)位置X3为A或P;

[0064] (d)位置X4为E、S、D或A;且

[0065] (e)位置X5为V、L或E;和

[0066] (ii)[N3]包含VQK、VQN、EQK、VKK、VHK、VQQ或LQK的氨基酸序列;或

[0067] 其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0068] 2.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包括包含下式的胺基序列:

[0069] [N2]-[N3],其中:

[0070] (i)[N2]包含位置X1、X2、X3、X4和X5,其中:

[0071] (a)位置X1为Y、N或C;

[0072] (b)位置X2为P、K、T或Q;

[0073] (c)位置X3为A或P;

[0074] (d)位置X4为E、S或A;且

[0075] (e)位置X5为V、L或E;和

[0076] (ii)[N3]包含VQK、EQK、VKK、VHK、VQQ或LQK的氨基酸序列;或

[0077] 其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,

保守取代)。

[0078] 3.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:

[0079] (i)[N0]包含TNN、TNT、INN、TNS、NNN或TNK;

[0080] (ii)[N1]包含QSS、QSK、TSL、SSS、QSR、AGA、IGS、QAS、ASS、LGS、QST、HSS、LSS或QRS;

[0081] (iii)[N2]包含YPAEV(SEQ ID NO:1)、YPPSL(SEQ ID NO:2)、NKAEV(SEQ ID NO:3)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、YQAEV(SEQ ID NO:6)、YTPSL(SEQ ID NO:7)、YPAAV(SEQ ID NO:8)、NPAEV(SEQ ID NO:9)、CPAEV(SEQ ID NO:10)或YQAEE(SEQ ID NO:11);

[0082] (iv)[N3]包含VQK、EQK、VKK、VHK、VQQ或LQK;和

[0083] (v)[N4]包含TA、PA或NA;且

[0084] 任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0085] 4.如实施方案1或2的AAV衣壳变体,其中:

[0086] (a)位置X1为Y或N;

[0087] (b)位置X2为P、T或Q;

[0088] (c)位置X3为A;

[0089] (d)位置X4为E或S;和/或

[0090] (e)位置X5为V或L。

[0091] 5.如实施方案1、2或4中任一项的AAV衣壳变体,其中[N2]包含YP、NK、YT、YQ、NP、CP、TH、AE、PS、AA、AS、PA、PP、KA、TA、QA、TP、HA、EV、SL、EE、AV或SH。

[0092] 6.如实施方案1、2、4或5中任一项的AAV衣壳变体,其中[N2]包含YPA、YPP、NKA、YTA、YQA、YTP、NPA、CPA、THA、PAE、PPS、KAE、TAE、QAE、TPS、PAA、HAS、AEV、PSL、AEE或AAV。

[0093] 7.如实施方案1-2或4-6中任一项的AAV衣壳变体,其中[N2]包含YPAE(SEQ ID NO:21)、YPPS(SEQ ID NO:22)、NKAE(SEQ ID NO:23)、YTAE(SEQ ID NO:24)、YQAE(SEQ ID NO:25)、YTPS(SEQ ID NO:26)、YPAA(SEQ ID NO:27)、NPAE(SEQ ID NO:28)、CPAE(SEQ ID NO:29)、THAS(SEQ ID NO:30)、PAEV(SEQ ID NO:17)、PPSL(SEQ ID NO:31)、KAEV(SEQ ID NO:32)、TAEV(SEQ ID NO:16)、PAEE(SEQ ID NO:18)、QAEV(SEQ ID NO:15)、TPSL(SEQ ID NO:33)、PAAV(SEQ ID NO:34)或QAEE(SEQ ID NO:35)。

[0094] 8.如实施方案1-7中任一项的AAV衣壳变体,其中[N2]为或包含YPAEV(SEQ ID NO:1)、YPPSL(SEQ ID NO:2)、NKAEV(SEQ ID NO:3)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、YQAEV(SEQ ID NO:6)、YTPSL(SEQ ID NO:7)、YPAAV(SEQ ID NO:8)、NPAEV(SEQ ID NO:9)、CPAEV(SEQ ID NO:10)或YQAEE(SEQ ID NO:11)。

[0095] 9.如实施方案1-2或4-8中任一项的AAV衣壳变体,其中[N3]包含VQK、EQK或VKK的氨基酸序列。

[0096] 10.如实施方案1-2或4-9中任一项的AAV衣壳变体,其中[N3]包含VQK。

[0097] 11.如实施方案1-2或4-9中任一项的AAV衣壳变体,其中[N3]包含EQK。

[0098] 12.如实施方案1-2或4-9中任一项的AAV衣壳变体,其中[N3]包含VKK

[0099] 13.如实施方案1-12中任一项的AAV衣壳变体,其中[N2]为或包含YPAEV(SEQ ID

NO:1)的氨基酸序列且[N3]为或包含VQK的氨基酸序列。

[0100] 14.如实施方案1-12中任一项的AAV衣壳变体,其中:

[0101] (i)[N2]为或包含YTPSL(SEQ ID NO:7)的氨基酸序列且[N3]为或包含VQK的氨基酸序列;

[0102] (ii)[N2]为或包含YPPSL(SEQ ID NO:2)的氨基酸序列且[N3]为或包含VQK的氨基酸序列;

[0103] (iii)[N2]为或包含YPPS(SEQ ID NO:2)L的氨基酸序列且[N3]为或包含EQK的氨基酸序列;或

[0104] (iv)[N2]为或包含YPPSL(SEQ ID NO:2)的氨基酸序列且[N3]为或包含VKK的氨基酸序列。

[0105] 15.如实施方案1-2或4-14中任一项的AAV衣壳变体,其中[N2]-[N3]包含:

[0106] (i)AEVVQK(SEQ ID NO:36)、PSLVQK(SEQ ID NO:37)、AEVEQK(SEQ ID NO:38)、AEEVQK(SEQ ID NO:39)、PSLEQK(SEQ ID NO:40)、PSLVKK(SEQ ID NO:41)、AEVVKK(SEQ ID NO:42)、AEVVHK(SEQ ID NO:43)、AAVVQK(SEQ ID NO:44)、AEVVQQ(SEQ ID NO:45)或AEVLQK(SEQ ID NO:46);

[0107] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0108] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0109] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0110] 16.如实施方案1-2或4-15中任一项的AAV衣壳变体,其中[N2]-[N3]包含:

[0111] (i)PAEVVQK(SEQ ID NO:20)、PPSLVQK(SEQ ID NO:47)、KAEVVQK(SEQ ID NO:48)、TAEVVQK(SEQ ID NO:49)、PAEVEQK(SEQ ID NO:50)、PAEEVQK(SEQ ID NO:51)、QAEVVQK(SEQ ID NO:52)、TPSLVQK(SEQ ID NO:53)、PPSLEQK(SEQ ID NO:54)、PPSLVKK(SEQ ID NO:55)、PAEVVKK(SEQ ID NO:56)、PAEVVHK(SEQ ID NO:57)、PAAVVQK(SEQ ID NO:58)、PAEVVQQ(SEQ ID NO:59)、TAEVVKK(SEQ ID NO:60)、PAEVLQK(SEQ ID NO:61)或QAEEVQK(SEQ ID NO:62);

[0112] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;

[0113] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0114] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0115] 17.如实施方案1-16中任一项的AAV衣壳变体,其中[N2]-[N3]为或包含:

[0116] (i)YPAEVVQK(SEQ ID NO:943)、YPPSLVQK(SEQ ID NO:946)、NKAEVVQK(SEQ ID NO:947)、YTAEVVQK(SEQ ID NO:948)、YPAEVEQK(SEQ ID NO:949)、YPAEEVQK(SEQ ID NO:950)、YQAEVVQK(SEQ ID NO:951)、YTPSLVQK(SEQ ID NO:952)、YPPSLEQK(SEQ ID NO:953)、YPPSLVKK(SEQ ID NO:954)、YPAEVVKK(SEQ ID NO:955)、YPAEVVHK(SEQ ID NO:956)、YPAAVVQK(SEQ ID NO:957)、NPAEVVQK(SEQ ID NO:958)、YPAEVVQQ(SEQ ID NO:959)、

CPAEVVQK(SEQ ID NO:960)、YTAEVVKK(SEQ ID NO:961)、YPAEVLQK(SEQ ID NO:962)或YQAEEVQK(SEQ ID NO:963)；

[0117] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列；

[0118] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0119] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0120] 18.如实施方案1-17中任一项的AAV衣壳变体,所述AAV衣壳变体进一步包含以下一个、两个、三个或全部:相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处非Q的氨基酸(例如,T、S、A、I、L或H),位置575处非S的氨基酸(例如,G、A或R),和/或位置576处非S的氨基酸(例如,K、L、R、A或T)。

[0121] 19.如实施方案1-17中任一项的AAV衣壳变体,所述AAV衣壳变体进一步包含：

[0122] (i)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的S和/或位置576处的S；

[0123] (ii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的T、位置575处的S和/或位置576处的L；

[0124] (iii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的S、位置575处的S和/或位置576处的S；

[0125] (iv)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的S和/或位置576处的R；

[0126] (v)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的S和/或位置576处的K；

[0127] (vi)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的A、位置575处的G和/或位置576处的A；

[0128] (vii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的I、位置575处的G和/或位置576处的S；

[0129] (viii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的A和/或位置576处的S；

[0130] (ix)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的A、位置575处的S和/或位置576处的S；

[0131] (x)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的L、位置575处的G和/或位置576处的S；

[0132] (xi)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的S和/或位置576处的T；

[0133] (xii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的H、位置575处的S和/或位置576处的S；

[0134] (xiii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的L、位置575处的S和/或位置576处的S；或

[0135] (xiv)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的R和/或位置576处的S。

[0136] 20.如实施方案1-19中任一项的AAV衣壳变体,其进一步包含[N1],其中[N1]包含位置X_D、X_E和X_F,其中:

[0137] (a)位置X_D为Q、T、S、A、I、L或H;

[0138] (b)位置X_E为S、G、A或R;且

[0139] (c)位置X_F为S、K、L、R、A或T;且

[0140] 任选地,其中所述AAV衣壳变体包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0141] 21.如实施方案20的AAV衣壳变体,其中[N1]包含SK、SL、SS、SR、GA、GS、AS、ST、RS、QS、TS、AG、IG、QA、LG、HS、LS或QR。

[0142] 22.如实施方案3、20或21中任一项的AAV衣壳变体,其中[N1]为或包含QSS、QSK、TSL、SSS、QSR、AGA、IGS、QAS、ASS、LGS、QST、HSS、LSS或QRS。

[0143] 23.如实施方案20-22中任一项的AAV衣壳变体,其中[N1]-[N2]包含:

[0144] (i)SSYPA(SEQ ID NO:63)、SKYPA(SEQ ID NO:64)、SLYPA(SEQ ID NO:65)、SRYPA(SEQ ID NO:66)、SSYPP(SEQ ID NO:67)、GAYPA(SEQ ID NO:68)、GSYPA(SEQ ID NO:69)、ASYPA(SEQ ID NO:70)、STNKA(SEQ ID NO:71)、SSYTA(SEQ ID NO:72)、SSYQA(SEQ ID NO:73)、SSYTP(SEQ ID NO:74)、SSNPA(SEQ ID NO:75)、SLCPA(SEQ ID NO:76)、RSYTA(SEQ ID NO:77)或SSTHA(SEQ ID NO:78);

[0145] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;

[0146] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0147] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0148] 24.如实施方案20-23中任一项的AAV衣壳变体,其中[N1]-[N2]包含:

[0149] (i)SSYPAE(SEQ ID NO:79)、SKYPAE(SEQ ID NO:80)、SLYPAE(SEQ ID NO:81)、SRYPAE(SEQ ID NO:82)、SSYPPS(SEQ ID NO:83)、GAYPAE(SEQ ID NO:84)、GSYPAE(SEQ ID NO:85)、ASYPAE(SEQ ID NO:86)、STNKAE(SEQ ID NO:87)、SSYTAE(SEQ ID NO:88)、SSYQAE(SEQ ID NO:89)、SSYTPS(SEQ ID NO:90)、SSYPAA(SEQ ID NO:91)、SSNPAE(SEQ ID NO:92)、SLCPAE(SEQ ID NO:93)、RSYTAE(SEQ ID NO:94)、SSTHAS(SEQ ID NO:95);

[0150] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0151] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0152] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0153] 25.如实施方案3或20-24中任一项的AAV衣壳变体,其中[N1]-[N2]为或包含:

[0154] (i)QSSYPAEV(SEQ ID NO:96)、QSKYPAEV(SEQ ID NO:97)、TSLYPAEV(SEQ ID NO:

98)、SSSYPAEV(SEQ ID NO:99)、QSRYPAEV(SEQ ID NO:100)、QSSYPPSL(SEQ ID NO:101)、AGAYPAEV(SEQ ID NO:102)、IGSYPAEV(SEQ ID NO:103)、QASYPAEV(SEQ ID NO:104)、ASSYPAEV(SEQ ID NO:105)、LGSYPAEV(SEQ ID NO:106)、QSTNKAEV(SEQ ID NO:107)、HSSYPAEV(SEQ ID NO:108)、SSSYTAEV(SEQ ID NO:109)、TSLYPAEE(SEQ ID NO:110)、ASSYQAEV(SEQ ID NO:111)、QSSYTPSL(SEQ ID NO:112)、QSRYPAEE(SEQ ID NO:113)、LSSYQAEV(SEQ ID NO:114)、HSSYPAAV(SEQ ID NO:115)、QSSNPAEV(SEQ ID NO:116)、QSSYTAEV(SEQ ID NO:117)、TSLCPAEV(SEQ ID NO:118)、QRSYTAEV(SEQ ID NO:119)或QSSYQAEE(SEQ ID NO:120)；

[0155] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列；

[0156] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0157] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0158] 26.如实施方案20-25中任一项的AAV衣壳变体,其中[N1]-[N2]-[N3]包含：

[0159] (i)SSYPAEVVQ(SEQ ID NO:121)、SKYPAEVVQ(SEQ ID NO:122)、SLYPAEVVQ(SEQ ID NO:123)、SRYPAEVVQ(SEQ ID NO:124)、SSYPPSLVQ(SEQ ID NO:125)、GAYPAEVVQ(SEQ ID NO:126)、GSYPAEVVQ(SEQ ID NO:127)、ASYPAEVVQ(SEQ ID NO:128)、STNKAEVVQ(SEQ ID NO:129)、SSYTAEVVQ(SEQ ID NO:130)、SKYPAEVEQ(SEQ ID NO:131)、SLYPAEEVQ(SEQ ID NO:132)、SSYQAEVVQ(SEQ ID NO:133)、SSYTPSLVQ(SEQ ID NO:134)、SRYPAEEVQ(SEQ ID NO:135)、SSYPPSLEQ(SEQ ID NO:136)、SSYPPSLVK(SEQ ID NO:140)、SSYPAEVVK(SEQ ID NO:141)、SKYPAEVVH(SEQ ID NO:142)、SSYPAAVVQ(SEQ ID NO:143)、SSNPAEVVQ(SEQ ID NO:144)、SLCPAEVVQ(SEQ ID NO:145)、RSYTAEVVQ(SEQ ID NO:146)、SSYTAEVVK(SEQ ID NO:147)、SSYPAEVLQ(SEQ ID NO:148)或SSYQAEEVQ(SEQ ID NO:149)；

[0160] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7或8个氨基酸,例如连续氨基酸)的氨基酸序列；

[0161] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0162] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0163] 27.如实施方案3或20-26中任一项的AAV衣壳变体,其中[N1]-[N2]-[N3]为或包含：

[0164] (i)QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、QSSYPPSLVQK(SEQ ID NO:155)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)、QSTNKAEVVQK(SEQ ID NO:161)、HSSYPAEVVQK(SEQ ID NO:162)、SSSYTAEVVQK(SEQ ID NO:163)、QSKYPAEVEQK(SEQ ID NO:164)、TSLYPAEEVQK(SEQ ID NO:165)、ASSYQAEVVQK(SEQ ID NO:166)、QSSYTPSLVQK(SEQ ID NO:167)、QSRYPAEEVQK(SEQ ID NO:168)、

QSSYPPSLEQK(SEQ ID NO:169)、QSSYPPSLVKK(SEQ ID NO:170)、LSSYQAEVVQK(SEQ ID NO:171)、SSSYPAEVVKK(SEQ ID NO:172)、QSKYPAEVVHK(SEQ ID NO:173)、HSSYPAAVVQK(SEQ ID NO:174)、QSSNPAEVVQK(SEQ ID NO:175)、SSSYPAEVVQQ(SEQ ID NO:176)、QSSYTAEVVQK(SEQ ID NO:177)、TSLCPAEVVQK(SEQ ID NO:178)、QRSYTAEVVQK(SEQ ID NO:179)、QSSYTAEVVKK(SEQ ID NO:180)、HSSYPAEVLQK(SEQ ID NO:181)或QSSYQAEEVQK(SEQ ID NO:182)；

[0165] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列；

[0166] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0167] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0168] 28.如实施方案1-27中任一项的AAV衣壳变体,其进一步包含[N0],其中[N0]包含位置X_A、X_B和X_C,其中：

[0169] (a)位置X_A为T、I或N；

[0170] (b)位置X_B为N；

[0171] (c)位置X_C为N、T、S或K；且

[0172] 任选地,其中所述AAV衣壳变体包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0173] 29.如实施方案28的AAV衣壳变体,其中[N0]包含TN、IN、NN、NT、NS或NK。

[0174] 30.如实施方案3、28或29的AAV衣壳变体,其中[N0]为或包含TNN、TNT、INN、TNS、NNN或TNK。

[0175] 31.如实施方案3或28-30中任一项的AAV衣壳变体,其中[N0]-[N1]为或包含：

[0176] (i)TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNQST(SEQ ID NO:193)、TNNHSS(SEQ ID NO:194)、TNNQSK(SEQ ID NO:184)、TNNLSS(SEQ ID NO:195)、INNQSS(SEQ ID NO:196)、TNSQSS(SEQ ID NO:197)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)、TNNQRS(SEQ ID NO:200)或TNKQAS(SEQ ID NO:201)；

[0177] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列；

[0178] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0179] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0180] 32.如实施方案3或28-31中任一项的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]为或包含：

[0181] (i)TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAE VVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVV QK(SEQ ID NO:510)、TNNQSSYPPSLVQK(SEQ ID NO:512)、TNNAGAYPAEVVQK(SEQ ID NO:513)、

TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQK(SEQ ID NO:517)、TNT ASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNQSTNKAEVVQK(SEQ ID NO:524)、TNNHSS YPAEVVQK(SEQ ID NO:525)、TNNSSSYTAEVVQK(SEQ IDNO:526)、TNNQSKYPAEVEQK(SEQ ID NO:529)、TNNTSLYP AEEVQK(SEQ ID NO:530)、TNTASSYQAEVVQK(SEQ ID NO:531)、TNNQSSYTPSLVQK(SEQ ID NO:533)、TNNQSRYPAEE VQK(SEQ ID NO:534)、TNNQSSYPPSLEQK(SEQ ID NO:535)、TNNQSSYPPSLVKK(SEQ ID NO:536)、TNNLSSYQAEVVQ K(SEQ ID NO:539)、TNNSSSYPAEVVKK(SEQ ID NO:540)、TNNQSKYPAEVVHK(SEQ ID NO:542)、INNQSSYPAEVVQK(SEQ ID NO:543)、TNNHSSYPAAVVQK(SEQ ID NO:545)、TNS QSSNPAEVVQK(SEQ ID NO:548)、TNNSSSYPAEVVQQ(SEQ ID NO:551)、NNNQSRYPAEVVQK(SEQ ID NO:552)、TNNQSS YTAEVVQK(SEQ ID NO:553)、TNNTSLCPAEVVQK(SEQ ID NO:554)、TNSTSLYPAEVVQK(SEQ ID NO:556)、TNNQRSYT AEVVQK(SEQ ID NO:557)、TNNQSSYTAEVVKK(SEQ ID NO:558)、TNNHSSYPAEVLQK(SEQ ID NO:560)、TNNQSSYQAEE VQK(SEQ ID NO:562)或TNKQASYPAEVVQK(SEQ ID NO:563)；

[0182] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列；

[0183] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0184] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0185] 33.如实施方案3或32的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)。

[0186] 34.如实施方案3或32的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]为或包含TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNTSL YPAEVVQK(SEQ ID NO:506)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNQSSYTPSLVQK(SEQ ID NO:533)、TNNQSSYPP SLVQK(SEQ ID NO:512)、TNNQSRYPAEVVQK(SEQ ID NO:510)、TNNQSSYPPSLEQK(SEQ ID NO:535)、TNNQSSYPPSLVK K(SEQ ID NO:536)或INNQSSYPAEVVQK(SEQ ID NO:543)。

[0187] 35.如实施方案1-34中任一项的AAV衣壳变体,其进一步包含[N4],其中[N4]包含位置X_G和X_H,其中：

[0188] (a)位置X_G为T、P或N;且

[0189] (b)位置X_H为A;且

[0190] 任选地,其中所述AAV衣壳变体包含(a)或(b)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0191] 36.如实施方案35的AAV衣壳变体,其中[N4]为或包含TA、PA或NA。

[0192] 37.如实施方案3、35或36中任一项的AAV衣壳变体,其中[N3]-[N4]为或包含：

[0193] (i)VQKTA(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKKTA(SEQ ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SEQ ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571)；

[0194] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列；

[0195] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0196] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0197] 38.如实施方案3或35-37中任一项的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]-[N4]为或包含:

[0198] (i)TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNNQSKYP AEVVQKTA(SEQ ID NO:1538)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEVVQKTA(SEQ ID NO:1539)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)、TNNQSSYPPSLVQKT A(SEQ ID NO:1300)、TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNIGSYPAEVVQKTA(SEQ ID NO:1112)、TNNQASYPAEVVQKTA(SEQ ID NO:1194)、TNTASSYPAEVVQKTA(SEQ ID NO:1575)、TNNLGSYPAEVVQKTA(SEQ ID NO:1027)、TNN QSTNKAEVVQKTA(SEQ ID NO:1578)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TNNSSSYTAEVVQKTA(SEQ ID NO:1214)、TNNQSKYPAE VEQKTA(SEQ ID NO:1254)、TNNTSLYPAEEVQKTA(SEQ ID NO:1583)、TNTASSYQAEVVQKTA(SEQ ID NO:1584)、TNNQS SYTPSLVQKTA(SEQ ID NO:1585)、TNNQSRYPAEEVQKTA(SEQ ID NO:1342)、TNNQSSYPPSLEQKTA(SEQ ID NO:1590)、TNNQSSYPPSLVKKTA(SEQ ID NO:1591)、TNNLSSYQAEVVQK TA(SEQ ID NO:1592)、TNNQSSYPPSLVQKPA(SEQ ID NO:1593)、TNNSSSYPAEVVKKTA(SEQ ID NO:1331)、TNNQSKYPAEVVHKTA(SEQ ID NO:1453)、TNNSSSYPAEVVQKPA(SEQ ID NO:1142)、INNQSSYPAEVVQKTA(SEQ ID NO:1024)、TNN HSSYPAAVVQKTA(SEQ ID NO:1598)、TNSQSSNPAEVVQKTA(SEQ ID NO:1599)、TNNSSSYPAEVVQQTA(SEQ ID NO:1419)、NNNQSRYPAEVVQKTA(SEQ ID NO:1601)、TNNQSSYTAE VVQKNA(SEQ ID NO:1602)、TNNTSLCPAEVVQKTA(SEQ ID NO:1603)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)、TNNQR SYTAEVVQKTA(SEQ ID NO:1604)、TNNQSSYTAEVVKKTA(SEQ ID NO:1606)、TNNHSSYPAEVLQKTA(SEQ ID NO:1607)、TNNQSSYQAEEVQKTA(SEQ ID NO:1608)、TNKQASYPAEVVQ KTA(SEQ ID NO:1587);

[0199] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;

[0200] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0201] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0202] 39.如实施方案3或35-38的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]-[N4]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[0203] 40.如实施方案3或35-38的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]-[N4]为或包含TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNQSKYPAE VVQKTA(SEQ ID NO:1538)、TNNQSSYTPSLVQKTA(SEQ ID NO:1585)、TNNQSSYPPSLVQKTA(SEQ ID NO:1300)、TNNQS RYPAEVVQKTA(SEQ ID NO:1327)、TNNQSSYPPSLEQKTA(SEQ ID NO:1590)、TNNQSSYPPSLVKKTA(SEQ ID NO:1591)或INNQSSYPAEVVQKTA(SEQ ID NO:1024)。

[0204] 41.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包括包含下式的胺基序列:[B]-[C],其中

[0205] (i)[B]包含位置X1、X2和X3,其中:

[0206] (a)位置X1为Q、T、S、A、I、L或H;

[0207] (b)位置X2为S、G或A;且

[0208] (c)位置X3为S、K、L、R或A;和

[0209] (ii)[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;且

[0210] 任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0211] 42.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:

[0212] (i)[A],其中[A]包含TNN、TNT、INN、NNN、TNS或TNK的氨基酸序列;

[0213] (ii)[B],其中[B]包含QSS、TSL、SSS、QSR、QSK、AGA、IGS、QAS、ASS、LGS或HSS的氨基酸序列;

[0214] (iii)[C],其中[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;和

[0215] (iv)[D],其中[D]包含TA或PA的氨基酸序列;且

[0216] 任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代);且

[0217] 进一步任选地,其中[C]置换相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列的位置577。

[0218] 43.如实施方案41的AAV衣壳变体,其中:

[0219] (a)位置X1为Q、T、S、A或H;

[0220] (b)位置X2为S或G;且

[0221] (c)位置X3为S、K、L或R。

[0222] 44.如实施方案41或43的AAV衣壳变体,其中[B]包含QS、TS、SS、AG、IG、QA、AS、LG、HS、SK、SL、SR、GA或GS。

[0223] 45.如实施方案41-44中任一项的AAV衣壳变体,其中[B]为或包含QSS、TSL、SSS、QSR、QSK、AGA、IGS、QAS、ASS、LGS或HSS。

[0224] 46.如实施方案41或43-45中任一项的AAV衣壳变体,其中[B]-[C]包含:

[0225] (i)SSYPAEVVQK(SEQ ID NO:572)、SKYPAEVVQK(SEQ ID NO:573)、SLYPAEVVQK(SEQ ID NO:574)、SRYPAEVVQK(SEQ ID NO:575)、GAYPAEVVQK(SEQ ID NO:576)、GSYPAEVVQK(SEQ ID NO:580)或ASYPAEVVQK(SEQ ID NO:582);

[0226] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;

[0227] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0228] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0229] 47.如实施方案41-46中任一项的AAV衣壳变体,其中[B]-[C]为或包含:

[0230] (i)QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)或HSSYPAEVVQK(SEQ ID NO:162)；

[0231] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列；

[0232] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0233] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0234] 48.如实施方案41-47中任一项的AAV衣壳变体,其中[B]-[C]为或包含QSSYPAEVVQK(SEQ ID NO:150)。

[0235] 49.如实施方案41-47中任一项的AAV衣壳变体,其中[B]-[C]为或包含AGAYPAEVVQK(SEQ ID NO:156)、TSLYPAEVVQK(SEQ ID NO:152)、QSKYPAEVVQK(SEQ ID NO:151)或QSRYPAEVVQK(SEQ ID NO:154)。

[0236] 50.如实施方案1-49中任一项的AAV衣壳变体,其进一步包含以下一个或两个:相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处非T的氨基酸(例如,I或N),和/或位置573处非N的氨基酸(例如,T、S或K)。

[0237] 51.如实施方案1-49中任一项的AAV衣壳变体,其进一步包含:

[0238] (i)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的N;

[0239] (ii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的T;

[0240] (iii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处的I、位置572处的N和/或位置573处的N;

[0241] (iv)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的S;

[0242] (v)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处的N、位置572处的N和/或位置573处的N;或

[0243] (vi)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的K。

[0244] 52.如实施方案41-51中任一项的AAV衣壳变体,其进一步包含[A],其中[A]包含位置X_A、X_B及X_C,其中:

[0245] (a)位置X_A为T、I或N;

[0246] (b)位置X_B为N;且

[0247] (c)位置X_C为N、T、S或K;且

[0248] 任选地,其中所述AAV衣壳变体包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0249] 53.如实施方案52的AAV衣壳变体,其中[A]包含TN、IN、NN、NT、NS或NK。

[0250] 54.如实施方案42、52或53中任一项的AAV衣壳变体,其中[A]为或包含TNN、TNT、INN、NNN、TNS或TNK。

[0251] 55.如实施方案42或52-54中任一项的AAV衣壳变体,其中[A]-[B]为或包含:

[0252] (i)TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNHSS(SEQ ID NO:194)、INNQSS(SEQ ID NO:196)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)或TNKQAS(SEQ ID NO:201);

[0253] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0254] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0255] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0256] 56.如实施方案42或52-55中任一项的AAV衣壳变体,其中[A]-[B]-[C]为或包含:

[0257] (i)TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAE VVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVV QK(SEQ ID NO:510)、TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQ K(SEQ ID NO:517)、TNTASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNHSSYPAEVVQK(SEQ ID NO:525)、INNQSSYPAEVVQK(SEQ ID NO:543)、NNN QSRYPAEVVQK(SEQ ID NO:552)、TNSTSLYPAEVVQK(SEQ ID NO:556)或TNKQASYPAEVVQK(SEQ ID NO:563);

[0258] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;

[0259] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0260] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0261] 57.如实施方案42或52-54中任一项的AAV衣壳变体,其中[A]-[B]-[C]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)。

[0262] 58.如实施方案42或52-54中任一项的AAV衣壳变体,其中[A]-[B]-[C]为或包含TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNQSRYPAEVVQK(SEQ ID NO:510)或INNQSSYPAEVVQK(SEQ ID NO:543)。

[0263] 59.如实施方案1-58中任一项的AAV衣壳变体,其进一步包含:

[0264] (i)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处非T的氨基酸(例如,P或N);或

[0265] (ii)相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,位置585处非T的氨基酸(例如,P或N)。

[0266] 60.如实施方案1-58中任一项的AAV衣壳变体,其进一步包含:

[0267] (i)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的T和/或位置579处的A;或相对于根据SEQ ID NO:982编号的参考序列,位置585处的T和/或位置586处的A;

[0268] (ii)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的P和/或位置579处的A;或相对于根据SEQ ID NO:982编号的参考序列,位置585处的P和/或位置586处的A;或

[0269] (iii)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的N和/或位置579处的A;或相对于根据SEQ ID NO:982编号的参考序列,位置585处的N和/或位置586处的A。

[0270] 61.如实施方案41-60中任一项的AAV衣壳变体,其进一步包含[D],其中[D]包含位置X4和X5,其中:

[0271] (a)位置X4为T或N;且

[0272] (b)位置X5为A;且

[0273] 任选地,其中所述AAV衣壳变体包含(a)或(b)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0274] 62.如实施方案42或61的AAV衣壳变体,其中[D]为或包含TA或PA。

[0275] 63.如实施方案42、61或62中任一项的AAV衣壳变体,其中[C]-[D]为或包含:

[0276] (i)YPAEVVQKTA(SEQ ID NO:584)或YPAEVVQKPA(SEQ ID NO:586);

[0277] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;

[0278] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0279] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0280] 64.如实施方案42或61-63中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]为或包含:

[0281] (i)TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNNQSKYP AEVVQKTA(SEQ ID NO:1538)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEVVQKTA(SEQ ID NO:1539)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)、TNNAGAYPAEVVQK TA(SEQ ID NO:1021)、TNNIGSYPAEVVQKTA(SEQ ID NO:1112)、TNNQASYPAEVVQKTA(SEQ ID NO:1194)、TNTASSYPAEVVQKTA(SEQ ID NO:1575)、TNNLGSYPAEVVQKTA(SEQ ID NO:1027)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNN SSSYPAEVVQKPA(SEQ ID NO:1142)、INNQSSYPAEVVQKTA(SEQ ID NO:1024)、NNNQSRYPAEVVQKTA(SEQ ID NO:1601)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)或TNKQASYPAE VVQKTA(SEQ ID NO:1587);

[0282] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;

[0283] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0284] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0285] 65.如实施方案42或61-64中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[0286] 66.如实施方案42或61-64中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]为或包含TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNQSKYPAE VVQKTA(SEQ ID NO:1538)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)或INNQSSYPAEVVQKTA(SEQ ID NO:1024)。

[0287] 67.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包括包含下式的胺基序列:

[0288] [N2]-[N3],其中:

[0289] (i)[N2]包含位置X1、X2、X3、X4和X5,其中:

[0290] (a)位置X1为Y或T;

[0291] (b)位置X2为Q、T、P或E;

[0292] (c)位置X3为A;

[0293] (d)位置X4为E或D;且

[0294] (e)位置X5为V或E;和

[0295] (ii)[N3]包含VQK或VQN的氨基酸序列;或

[0296] 其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0297] 68.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:

[0298] (i)[N0]包含TNN、TNS、TNT或TNK;

[0299] (ii)[N1]包含QSS、SLS、SLY、SAT或QTS;

[0300] (iii)[N2]包含YPAEV(SEQ ID NO:1)、YQAEV(SEQ ID NO:6)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、TEAEV(SEQ ID NO:12)或YPADV(SEQ ID NO:13);

[0301] (iv)[N3]包含VQK或VQN;和

[0302] (v)[N4]包含TA、PA、TD、NA或PA;且

[0303] 任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0304] 69.如实施方案1或67的AAV衣壳变体,其中[N2]包含YP、YQ、YT、TE、QA、TA、PA、EA、EV、EE、DV、AE或AD。

[0305] 70.如实施方案1、67或69中任一项的AAV衣壳变体,其中[N2]包含YPA、YQA、YTA、TEA、QAE、TAE、PAE、EAE、PAD、AEV、AEE或ADV。

[0306] 71.如实施方案1、67、69或70中任一项的AAV衣壳变体,其中[N2]包含YPAE(SEQ ID NO:21)、YQAE(SEQ ID NO:25)、YTAE(SEQ ID NO:24)、TEAE(SEQ ID NO:587)、YPAD(SEQ ID NO:588)、QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)、EAEV(SEQ ID NO:590)或PADV(SEQ ID NO:19)。

[0307] 72.如实施方案1或67-71中任一项的AAV衣壳变体,其中[N2]为或包含YPAEV(SEQ

ID NO:1)、YQAEV(SEQ ID NO:6)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、TEAEV(SEQ ID NO:12)或YPADV(SEQ ID NO:13)。

[0308] 73.如实施方案1或67-72中任一项的AAV衣壳变体,其中[N3]包含VQK的氨基酸序列。

[0309] 74.如实施方案1、67或69-73中任一项的AAV衣壳变体,其中[N2]-[N3]包含:

[0310] (i)AEVVQK(SEQ ID NO:36)、AEEVQK(SEQ ID NO:39)、AEVVQN(SEQ ID NO:591)或ADVVQK(SEQ ID NO:593);

[0311] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0312] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0313] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0314] 75.如实施方案1、67或69-74中任一项的AAV衣壳变体,其中[N2]-[N3]包含:

[0315] (i)PAEVVQN(SEQ ID NO:594)、QAEVVQK(SEQ ID NO:52)、TAEVVQK(SEQ ID NO:49)、PAEVVQK(SEQ ID NO:20)、PAEEVQK(SEQ ID NO:51)、EAEVVQK(SEQ ID NO:595)或PADVVQK(SEQ ID NO:596);

[0316] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;

[0317] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0318] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0319] 76.如实施方案1或57-75中任一项的AAV衣壳变体,其中[N2]-[N3]为或包含:

[0320] (i)YPAEVVQK(SEQ ID NO:943)、YQAEVVQK(SEQ ID NO:951)、YTAEVVQK(SEQ ID NO:948)、YPAEEVQK(SEQ ID NO:950)、YPAEVVQN(SEQ ID NO:964)、TEAEVVQK(SEQ ID NO:965)或YPADVVQK(SEQ ID NO:966);

[0321] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;

[0322] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0323] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0324] 77.如实施方案1或67-76中任一项的AAV衣壳变体,其中[N2]-[N3]为或包含YPAEVVQK(SEQ ID NO:943)。

[0325] 78.如实施方案1-77中任一项的AAV衣壳变体,其中[N2]-[N3]置换相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列的位置577。

[0326] 79.如实施方案1或67-78中任一项的AAV衣壳变体,其进一步包含以下一个、两个、三个或全部:相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处非

Q的氨基酸(例如,S),位置575处非S的氨基酸(例如,L、A或T),和/或位置576处非S的氨基酸(例如,Y或T)。

[0327] 80.如实施方案1或67-78中任一项的AAV衣壳变体,其进一步包含:

[0328] (i)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的S和/或位置576处的S;

[0329] (ii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的S、位置575处的L和/或位置576处的S;

[0330] (iii)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的S、位置575处的L和/或位置576处的Y;

[0331] (iv)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的S、位置575处的A和/或位置576处的T;或

[0332] (v)相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,位置574处的Q、位置575处的T和/或位置576处的S。

[0333] 81.如实施方案1或67-80中任一项的AAV衣壳变体,其进一步包含[N1],其中[N1]包含位置X_D、X_E和X_F,其中:

[0334] (a)位置X_D为Q或S;

[0335] (b)位置X_E为S、L、A或T;且

[0336] (c)位置X_F为S、Y或T;且

[0337] 任选地,其中所述AAV衣壳变体包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0338] 82.如实施方案81的AAV衣壳变体,其中[N1]包含QS、SL、SA、QT、LS、LY、AT、TS或SS。

[0339] 83.如实施方案68、81、82中任一项的AAV衣壳变体,其中[N1]为或包含QSS、SLS、SLY、SAT或QTS。

[0340] 84.如实施方案81-83中任一项的AAV衣壳变体,其中[N1]-[N2]包含:

[0341] (i)SSYPA(SEQ ID NO:63)、LSYQA(SEQ ID NO:597)、LSYTA(SEQ ID NO:598)、LYYPA(SEQ ID NO:600)、ATYPA(SEQ ID NO:601)、LSYPA(SEQ ID NO:603)或TSTEA(SEQ ID NO:605);

[0342] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;

[0343] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0344] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0345] 85.如实施方案81-84中任一项的AAV衣壳变体,其中[N1]-[N2]包含:

[0346] (i)SSYPAE(SEQ ID NO:79)、LSYQAE(SEQ ID NO:607)、LSYTAE(SEQ ID NO:610)、LYYPAE(SEQ ID NO:611)、ATYPAE(SEQ ID NO:613)、LSYPAE(SEQ ID NO:616)、TSTEAE(SEQ ID NO:619)或LSYPAD(SEQ ID NO:621);

[0347] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0348] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0349] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0350] 86.如实施方案68或81-85中任一项的AAV衣壳变体,其中[N1]-[N2]为或包含:

[0351] (i)QSSYPAEV(SEQ ID NO:96)、SLSYQAEV(SEQ ID NO:622)、SLSYTAEV(SEQ ID NO:623)、SLYYPAEV(SEQ ID NO:624)、SATYPAEV(SEQ ID NO:625)、SLSYPAEV(SEQ ID NO:629)、SLSYPAEE(SEQ ID NO:632)、QTSTEAEV(SEQ ID NO:633)或SLSYPADV(SEQ ID NO:634);

[0352] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;

[0353] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0354] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0355] 87.如实施方案68或81-86中任一项的AAV衣壳变体,其中[N1]-[N2]-[N3]为或包含:

[0356] (i)QSSYPAEVVQK(SEQ ID NO:150)、SLSYQAEVVQK(SEQ ID NO:635)、SLSYTAEVVQK(SEQ ID NO:637)、SLYYPAEVVQK(SEQ ID NO:639)、SATYPAEVVQK(SEQ ID NO:641)、SLSYPAEVVQK(SEQ ID NO:642)、SLSYPAEEVQK(SEQ ID NO:643)、SLSYPAEVVQN(SEQ ID NO:644)、QTSTEAEVVQK(SEQ IDNO:645)或SLSYPADVVQK(SEQ ID NO:646);

[0357] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11或12个氨基酸,例如连续氨基酸)的氨基酸序列;

[0358] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0359] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0360] 88.如实施方案68或81-87中任一项的AAV衣壳变体,其中[N1]-[N2]-[N3]为或包含QSSYPAEVVQK(SEQ ID NO:150)。

[0361] 89.如实施方案1或67-88中任一项的AAV衣壳变体,其进一步包含[NO],其中[NO]包含位置X_A、X_B和X_C,其中:

[0362] (a)位置X_A为T;

[0363] (b)位置X_B为N;且

[0364] (c)位置X_C为N、T、S或K;且

[0365] 任选地,其中所述AAV衣壳包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0366] 90.如实施方案89的AAV衣壳变体,其中[NO]:

[0367] (i)包含TN、NS、NT、NN或NK;和/或

[0368] (ii)为或包含TNS、TNT、TNN或TNK。

[0369] 91.如实施方案68、89或90中任一项的AAV衣壳变体,其中[NO]-[N1]为或包含:

[0370] (i)TNNQSS(SEQ ID NO:183)、TNSSLS(SEQ ID NO:647)、TNSSLY(SEQ ID NO:648)、TNTSAT(SEQ ID NO:649)、TNNQTS(SEQ ID NO:650)或TNKSAT(SEQ ID NO:651;

[0371] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0372] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0373] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0374] 92.如实施方案68或89-91中任一项的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]为或包含:

[0375] (i)TNNQSSYPAEVVQK(SEQ ID NO:500)、TNSSLSYQAEV VQK(SEQ ID NO:652)、TNSSLSYTAEVVQK(SEQ ID NO:654)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNTSATYPAEVVQ K(SEQ ID NO:656)、TNSSLSYPAEVVQK(SEQ ID NO:657)、TNSSLSYPAEEVQK(SEQ ID NO:658)、TNSSLSYPAEVVQN(SE Q ID NO:660)、TNNQTSTEAEVVQK(SEQ ID NO:662)、TNKS ATYPAEVVQK(SEQ ID NO:663)或TNSSLSYPADVVQK(SEQ ID NO:665);

[0376] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;

[0377] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0378] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0379] 93.如实施方案68或89-92中任一项的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)。

[0380] 94.如实施方案1或67-93中任一项的AAV衣壳变体,其进一步包含[N4],其中[N4]包含位置X_G和X_H,其中:

[0381] (a)位置X_G为T、P或N;且

[0382] (b)位置X_H为A或D;且

[0383] 任选地,其中所述AAV衣壳变体包含(a)或(b)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0384] 95.如实施方案94的AAV衣壳变体,其中[N4]为或包含TA、TD、PA或NA。

[0385] 96.如实施方案68、94或95的AAV衣壳变体,其中[N3]-[N4]为或包含:

[0386] (i)VQKTA(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKKTA(SEQ ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SEQ ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571);

[0387] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;

[0388] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0389] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨

基酸的氨基酸序列。

[0390] 97.如实施方案68或94-96中任一项的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]-[N4]为或包含:

[0391] (i)TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNSSLSYQ AEVVQKTA(SEQ ID NO:2064)、TNSSLSYTAEVVQKTA(SEQ ID NO:2065)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNTSATYPAEVVQKTA(SEQ ID NO:2067)、TNSSLSYPAEVVQKTA(SEQ ID NO:2068)、TNSSLSYPAEEVQKTA(SEQ ID NO:2069)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQN TA(SEQ ID NO:2071)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)、TNNQTSTEA EVVQKTA(SEQ ID NO:2074)、TNKSATYPAEVVQKTA(SEQ ID NO:2075)或TNSSLSYPADVVQKTA(SEQ ID NO:2076);

[0392] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;

[0393] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0394] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0395] 98.如实施方案68或94-97中任一项的AAV衣壳变体,其中[N0]-[N1]-[N2]-[N3]-[N4]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[0396] 99.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含式[K1]-[K2],其中:

[0397] (i)[K1]包含LSY或LYY;和

[0398] (ii)[K2]包含位置X1、X2、X3和X4,其中:

[0399] (a)位置X1为Q、T或P;

[0400] (b)位置X2为A;

[0401] (c)位置X3为E或D;且

[0402] (d)位置X4为V或E;且

[0403] 任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0404] 100.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:

[0405] (i)[K0],其包含TNNS(SEQ ID NO:14);

[0406] (ii)[K1],其包含LSY或LYY;

[0407] (iii)[K2],其包含QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)或PADV(SEQ ID NO:19);

[0408] (iv)[K3],其包含VQK或VQN;和

[0409] (v)[K4],其包含TA、TD、NA或PA;且

[0410] 任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0411] 101.如实施方案99或100的AAV衣壳变体,其中[K1]包含LSY。

[0412] 102.如实施方案99或101的AAV衣壳变体,其中[K2]包含QA、TA、PA、EV、EE、DV、AE或AD。

[0413] 103.如实施方案99、101或102中任一项的AAV衣壳变体,其中[K2]包含QAE、TAE、PAE、PAD、AEV、AEE或ADV。

[0414] 104.如实施方案99-103中任一项的AAV衣壳变体,其中[K2]为或包含QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)或PADV(SEQ ID NO:19)。

[0415] 105.如实施方案99或101-104中任一项的AAV衣壳变体,其中[K1]-[K2]包含LSYQA(SEQ ID NO:597)、LSYTA(SEQ ID NO:598)、LYYPA(SEQ ID NO:600)或LSYPA(SEQ ID NO:603)。

[0416] 106.如实施方案99或101-105中任一项的AAV衣壳变体,其中[K1]-[K2]包含:

[0417] (i)LSYQAE(SEQ ID NO:607)、LSYTAE(SEQ ID NO:610)、LYYPAE(SEQ ID NO:611)、LSYPAE(SEQ ID NO:616)或LSYPAD(SEQ ID NO:621);

[0418] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0419] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0420] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0421] 107.如实施方案99-106中任一项的AAV衣壳变体,其中[K1]-[K2]为或包含:

[0422] (i)LSYQAEV(SEQ ID NO:667)、LSYTAEV(SEQ ID NO:668)、LYYPAEV(SEQ ID NO:669)、LSYPAEV(SEQ ID NO:671)、LSYPAEE(SEQ ID NO:673)或LSYPADV(SEQ ID NO:674);

[0423] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;

[0424] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0425] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0426] 108.如实施方案99-107中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述AAV衣壳变体进一步包含位置574处非Q的氨基酸(例如,S)。

[0427] 109.如实施方案99-108中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述AAV衣壳变体进一步包含位置574处的S。

[0428] 110.如实施方案99-109中任一项的AAV衣壳变体,其进一步包含[K0],其中[K0]为或包含TNNS(SEQ ID NO:14);包含氨基酸序列的任何部分(例如,任何2或3个氨基酸,例如连续氨基酸)的氨基酸序列;相对于TNNS(SEQ ID NO:14)的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于TNNS(SEQ ID NO:14)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0429] 111.如实施方案110的AAV衣壳变体,其中[K0]-[K1]包含:

[0430] (i)TNSSLS(SEQ ID NO:647)或TNSSLY(SEQ ID NO:648)；

[0431] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列；

[0432] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0433] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0434] 112.如实施方案100、110或111中任一项的AAV衣壳变体,其中[K0]-[K1]为或包含：

[0435] (i)TNSSLSY(SEQ ID NO:676)或TNSSLYY(SEQ ID NO:678)；

[0436] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列；

[0437] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0438] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0439] 113.如实施方案110-112中任一项的AAV衣壳变体,其中[K0]-[K1]-[K2]包含：

[0440] (i)TNSSLSYQA(SEQ ID NO:679)、TNSSLSYTA(SEQ ID NO:681)、TNSSLYYPA(SEQ ID NO:682)或TNSSLSYPA(SEQ ID NO:683)；

[0441] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7或8个氨基酸,例如连续氨基酸)的氨基酸序列；

[0442] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0443] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0444] 114.如实施方案110-113中任一项的AAV衣壳变体,其中[K0]-[K1]-[K2]包含：

[0445] (i)TNSSLSYQAE(SEQ ID NO:684)、TNSSLSYTAE(SEQ ID NO:685)、TNSSLYYPAE(SEQ ID NO:686)、TNSSLSYPAE(SEQ ID NO:687)或TNSSLSYPAD(SEQ ID NO:689)；

[0446] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列；

[0447] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0448] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0449] 115.如实施方案100或110-114中任一项的AAV衣壳变体,其中[K0]-[K1]-[K2]为或包含：

[0450] (i)TNSSLSYQAEV(SEQ ID NO:692)、TNSSLSYTAEV(SEQ ID NO:693)、TNSSLYYPAEV(SEQ ID NO:696)、TNSSLSYPAEV(SEQ ID NO:697)、TNSSLSYPAEE(SEQ ID NO:698)或TNSSLSYPADV(SEQ ID NO:699)；

[0451] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;

[0452] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0453] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0454] 116.如实施方案99-115中任一项的AAV衣壳变体,其进一步包含[K3],其中[K3]包含位置X_A、X_B和X_C,其中:

[0455] (a)位置X_A为V;

[0456] (b)位置X_B为Q;且

[0457] (c)位置X_C为K或N;且

[0458] 任选地,其中所述AAV衣壳包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0459] 117.如实施方案116的AAV衣壳变体,其中[K3]:

[0460] (i)包含VQ、QK或QN;

[0461] (ii)为或包含VQK或VQN。

[0462] 118.如实施方案100、116或117的AAV衣壳变体,其中[K2]-[K3]为或包含:

[0463] (i)QAEVVQK(SEQ ID NO:52)、TAEVVQK(SEQ ID NO:49)、PAEVVQK(SEQ ID NO:20)、PAEEVQK(SEQ ID NO:51)、PAEVVQN(SEQ ID NO:594)或PADVVQK(SEQ ID NO:596);

[0464] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;

[0465] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0466] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0467] 119.如实施方案100或116-118的AAV衣壳变体,其中[K1]-[K2]-[K3]为或包含:

[0468] (i)LSYQAEVVQK(SEQ ID NO:700)、LSYTAEVVQK(SEQ ID NO:701)、LYYPAEVVQK(SEQ ID NO:702)、LSYPAEVVQK(SEQ ID NO:703)、LSYPAEEVQK(SEQ ID NO:704)、LSYPAEVVQN(SEQ ID NO:706)或LSYPADVVQK(SEQ ID NO:708);

[0469] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;

[0470] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0471] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0472] 120.如实施方案100或116-119的AAV衣壳变体,其中[K0]-[K1]-[K2]-[K3]为或包含:

[0473] (i)TNSSLSYQAEVVQK(SEQ ID NO:652)、TNSSLSYTAEV VQK(SEQ ID NO:654)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNSSLSYPAEVVQK(SEQ ID NO:657)、TNSSLSYPAEEVQK

(SEQ ID NO:658)、TNSSLSYPAEVVQN(SEQ ID NO:660)或TNSSLSYPADVVQK(SEQ ID NO:665)；

[0474] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列；

[0475] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0476] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0477] 121.如实施方案99-120中任一项的AAV衣壳变体,其进一步包含[K4],其中[K4]包含位置X_D和X_E,其中：

[0478] (a)位置X_D为T、P或N;且

[0479] (b)位置X_E为A或D;且

[0480] 任选地,其中所述AAV衣壳变体包含(a)或(b)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0481] 122.如实施方案100或121的AAV衣壳变体,其中[K4]为或包含TA、TD、PA或NA。

[0482] 123.如实施方案100、121或122中任一项的AAV衣壳变体,其中[K3]-[K4]为或包含：

[0483] (i)VQKTA(SEQ ID NO:564)、VQKTD(SEQ ID NO:714)、VQNTA(SEQ ID NO:715)、VQKNA(SEQ ID NO:570)或VQKPA(SEQ ID NO:567)；

[0484] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列；

[0485] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0486] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0487] 124.如实施方案100或121-123中任一项的AAV衣壳变体,其中[K0]-[K1]-[K2]-[K3]-[K4]为或包含：

[0488] (i)TNSSLSYQAEVVQKTA(SEQ ID NO:2064)、TNSSLSYT AEVVQKTA(SEQ ID NO:2065)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNSSLSYPAEVVQKTA(SEQ ID NO:2068)、TNSSLSYPAEEVQKTA(SEQ ID NO:2069)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQNTA(SEQ ID NO:2071)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQK PA(SEQ ID NO:2073)或TNSSLSYPADVVQKTA(SEQ ID NO:2076)；

[0489] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列；

[0490] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0491] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0492] 125.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含[B]-[C],其中：

[0493] (i)[B]包含位置X1、X2和X3,其中:

[0494] (a)位置X1为Q或S;

[0495] (b)位置X2为S、L或A;且

[0496] (c)位置X3为S、Y或T;和

[0497] (ii)[C]包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列;且

[0498] 任选地,其中所述AAV衣壳变体包含(i)和/或(ii)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0499] 126.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含以下一个、两个、三个、四个或全部:

[0500] (i)[A]包含TNN、TNS、TNT或TNK;

[0501] (ii)[B]包含QSS、SLY、SAT或SLS;

[0502] (iii)[C]包含YPAEVVQK(SEQ ID NO:943);和

[0503] (iv)[D]包含TA、TD、NA或PA;且

[0504] 任选地,其中所述AAV衣壳变体包含(i)-(v)中任何上述氨基酸的氨基酸修饰(例如,保守取代),且

[0505] 进一步任选地,其中[C]置换相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列的位置577。

[0506] 127.如实施方案125的AAV衣壳变体,其包含其中[B]包含QS、SL、SA、LY、AT、LS或SS。

[0507] 128.如实施方案125-127中任一项的AAV衣壳变体,其包含其中[B]为或包含QSS、SLY、SAT或SLS。

[0508] 129.如实施方案125、127或128中任一项的AAV衣壳变体,其中[B]-[C]包含:

[0509] (i)SSYPAEVVQK(SEQ ID NO:572)、LYYPAEVVQK(SEQ ID NO:702)、ATYPAEVVQK(SEQ ID NO:718)或LSYPAEVVQK(SEQ ID NO:703);

[0510] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;

[0511] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0512] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0513] 130.如实施方案125-129中任一项的AAV衣壳变体,其中[B]-[C]为或包含:

[0514] (i)QSSYPAEVVQK(SEQ ID NO:150)、SLYYPAEVVQK(SEQ ID NO:639)、SATYPAEVVQK(SEQ ID NO:641)或SLSYPAEVVQK(SEQ ID NO:642);

[0515] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;

[0516] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0517] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0518] 131.如实施方案126或130的AAV衣壳变体,其中[B]-[C]为或包含QSSYPAEVVQK(SEQ ID NO:150)。

[0519] 132.如实施方案67-98或125-131中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述AAV衣壳变体进一步包含位置573处非N的氨基酸(例如,T、S或K)。

[0520] 133.如实施方案67-98或125-131中任一项的AAV衣壳变体,其进一步包含:

[0521] (i)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的N;

[0522] (ii)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的T;

[0523] (iii)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的S;或

[0524] (iv)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置571处的T、位置572处的N和/或位置573处的K。

[0525] 134.如实施方案125-133中任一项的AAV衣壳变体,其进一步包含[A],其中[A]包含位置X_A、X_B和X_C,其中:

[0526] (a)位置X_A为T;

[0527] (b)位置X_B为N;且

[0528] (c)位置X_C为N、T、S或K;且

[0529] 任选地,其中所述AAV衣壳变体包含(a)-(c)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0530] 135.如实施方案126或134的AAV衣壳变体,其中[A]:

[0531] (i)包含TN、NS、NT、NK或NN;

[0532] (ii)为或包含TNN、TNS、TNT或TNK。

[0533] 136.如实施方案126、134或135中任一项的AAV衣壳变体,其中[A]-[B]为或包含:

[0534] (i)TNNQSS(SEQ ID NO:183)、TNSSLY(SEQ ID NO:648)、TNTSAT(SEQ ID NO:649)、TNSSLS(SEQ ID NO:647)或TNKSAT(SEQ ID NO:651);

[0535] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;

[0536] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0537] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0538] 137.如实施方案126或134-136中任一项的AAV衣壳变体,其中[A]-[B]-[C]为或包含:

[0539] (i)TNNQSSYPAEVVQK(SEQ ID NO:500)、TNSSLYYPAEV VQK(SEQ ID NO:655)、TNTSATYPAEVVQK(SEQ ID NO:656)、TNSSLSYPAEVVQK(SEQ ID NO:657)或TNKSATYPAEVVQ K(SEQ ID NO:663);

[0540] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12

或13个氨基酸,例如连续氨基酸)的氨基酸序列;

[0541] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0542] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0543] 138.如实施方案126或134-137中任一项的AAV衣壳变体,其中[A]-[B]-[C]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)。

[0544] 139.如实施方案67-98或125-138中任一项的AAV衣壳变体,其进一步包含以下一个或两个:相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处非T的氨基酸(例如,P或N)和/或位置579处非A的氨基酸(例如,D)。

[0545] 140.如实施方案67-98或125-138中任一项的AAV衣壳变体,其进一步包含:

[0546] (i)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的T和/或位置579处的A;

[0547] (ii)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的T和/或位置579处的D;

[0548] (iii)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的P和/或位置579处的A;或

[0549] (iv)相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处的N和/或位置579处的A。

[0550] 141.如实施方案125-140中任一项的AAV衣壳变体,其进一步包含[D],其中[D]包含位置X4和X5,其中:

[0551] (a)位置X4为T、N或P;且

[0552] (b)位置X5为A或D;且

[0553] 任选地,其中所述AAV衣壳变体包含(a)或(b)中任何上述氨基酸的氨基酸修饰(例如,保守取代)。

[0554] 142.如实施方案141的AAV衣壳变体,其中[D]为或包含TA、TD、NA或PA。

[0555] 143.如实施方案126、141或142的AAV衣壳变体,其中[C]-[D]为或包含:

[0556] (i)YPAEVVQKTA(SEQ ID NO:584)、YPAEVVQKTD(SEQ ID NO:719)、YPAEVVQKNA(SEQ ID NO:724)或YPAEVVQKPA(SEQ ID NO:586);

[0557] (ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;

[0558] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或

[0559] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0560] 144.如实施方案126或141-143中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]为或包含:

[0561] (i)TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNSSLYYP AEVVQKTA(SEQ ID NO:2066)、TNTSATYPAEVVQKTA(SEQ ID NO:2067)、TNSSLSYPAEVVQKTA(SEQ ID NO:2068)、TNS

SLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)或TNKSATYPAEVVQKTA(SEQ ID NO:2075)；

[0562] (ii)(ii)包含(i)中的氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列；

[0563] (iii)相对于(i)中的任何氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列；或

[0564] (iv)相对于(i)中的任一氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0565] 145.如实施方案126或141-144中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[0566] 146.如实施方案1-40、59、60、67-98、139或140中任一项的AAV衣壳变体,其中[N2]-[N3]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0567] 147.如实施方案2、15-40、59、60、68、79-98、139、137或146中任一项的AAV衣壳变体,其中[NO]、[N1]和/或[N4]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0568] 148.如实施方案2、15-40、59、60、68、79-98、139、140、146或147中任一项的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]-[N4]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0569] 149.如实施方案1-40、59、60、67-98、139、140或146-148中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[N2]紧接在位置576之后存在。

[0570] 150.如实施方案1-40、59、60、67-98、139、140或146-149中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N2]置换位置577(例如,T577)。

[0571] 151.如实施方案1-40、59、60、67-98、139、140或146-150中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N2]紧接在位置576之后存在,且其中[N2]置换位置577(例如,T577)。

[0572] 152.如1-40、59、60、67-98、139、140或146-151中任一项的AAV衣壳变体,其中[N2]对应于SEQ ID NO:982的位置577-581(例如,Y577、P578、A579、E580、V581)。

[0573] 153.如实施方案1-152中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述AAV衣壳变体包含位置577处非T的氨基酸。

[0574] 154.如实施方案1-153中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,所述AAV衣壳变体包含位置577处的Y。

[0575] 155.如实施方案1-40、59、60、67-98、139、140或146-154中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[N2]的X1存在于位置

577(例如,T577)处,且[N2]的位置X2-X5紧接在位置577之后存在。

[0576] 156.如实施方案1-40、59、60、67-98、139、140或146-155中任一项的AAV衣壳变体,其中[N2]的X1对应于位置577(例如,Y577),[N2]的X2对应于位置578(例如,P588),[N2]的X3对应于位置579(例如,A579),[N2]的X4对应于位置580(例如,E580),且[N2]的X5对应于SEQ ID NO:982的位置581(例如,V581)。

[0577] 157.如实施方案1-40、59、60、67-98、139、140或146-156中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[N2]-[N3]紧接在位置576之后存在。

[0578] 158.如实施方案1-40、59、60、67-98、139、140或146-157中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N2]-[N3]置换位置577。

[0579] 159.如实施方案1-40、59、60、67-98、139、140或146-158中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N2]-[N3]紧接在位置576之后存在,且其中[N2]-[N3]置换位置577(例如,T577)。

[0580] 160.如实施方案1-40、59、60、67-98、139、140或146-159中任一项的AAV衣壳变体,其中[N2]-[N3]对应于SEQ ID NO:982的位置577-584(例如,Y577、P578、A579、E580、V581、V582、Q583、K584)。

[0581] 161.如实施方案1-40、59、60、67-98、139、140或146-160中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[N2]-[N3]-[N4]紧接在位置576之后存在。

[0582] 162.如实施方案1-40、59、60、67-98、139、140或146-161中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578和A579)。

[0583] 163.如实施方案1-40、59、60、67-98、139、140或146-162中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N2]-[N3]-[N4]紧接在位置576之后存在,且其中[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578和A579)。

[0584] 164.如实施方案1-40、59、60、67-98、139、140或146-163中任一项的AAV衣壳变体,其中[N2]-[N3]-[N4]对应于SEQ ID NO:982的位置577-586(例如,Y577、P578、A579、E580、V581、V582、Q583、K584、T585、A586)。

[0585] 165.如实施方案3、18-40、59、60、68、79-98、139、140或146-164中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[N1]紧接在位置573之后存在。

[0586] 166.如实施方案3、18-40、59、60、68、79-98、139、140或146-165中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N1]置换位置574-576(例如,Q574、S575和S576)。

[0587] 167.如实施方案3、18-40、59、60、68、79-98、139、140或146-166中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N1]紧接在位置573之后存在,且其中[N1]置换位置574-576(例如,Q574、S575和S576)。

[0588] 168.如实施方案3、18-40、59、60、68、79-98、139、140或146-167中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[N1]置换位置574-

576(例如,Q574、S575和S576)。

[0589] 169.如实施方案3、18-40、59、60、68、79-98、139、140或146-168中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[N1]紧接在位置573之后存在,且其中[N1]置换位置574-576(例如,Q574、S575和S576)。

[0590] 170.如实施方案3、18-40、59、60、68、79-98、139、140或146-169中任一项的AAV衣壳变体,其中[N1]对应于SEQ ID NO:982的位置574-576(例如,Q574、S575和S576)。

[0591] 171.如实施方案3、18-40、59、60、68、79-98、139、140或146-170中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[N1]-[N2]-[N3]-[N4]紧接在位置573之后存在。

[0592] 172.如实施方案3、18-40、59、60、68、79-98、139、140或146-171中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N1]-[N2]-[N3]-[N4]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579)。

[0593] 173.如实施方案3、18-40、59、60、68、79-98、139、140或146-171中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[N1]-[N2]-[N3]-[N4]紧接在位置573之后存在,且其中[N1]-[N2]-[N3]-[N4]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579)。

[0594] 174.如实施方案3、18-40、59、60、68、79-98、139、140或146-173中任一项的AAV衣壳变体,其中[N1]-[N2]-[N3]-[N4]对应于SEQ ID NO:982的位置574-586(例如,Q574、S575、S576、Y577、P578、A579、E580、V581、V582、Q583、K584、T585、A586)。

[0595] 175.如实施方案3、18-40、59、60、68、79-98、139、140或146-174中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[NO]紧接在位置570之后存在。

[0596] 176.如实施方案3、18-40、59、60、68、79-98、139、140或146-175中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[NO]置换位置571-573(例如,T571、N572和N573)。

[0597] 177.如实施方案3、18-40、59、60、68、79-98、139、140或146-176中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[NO]紧接在位置570之后存在,且其中[NO]置换位置571-573(例如,T571、N572和N573)。

[0598] 178.如实施方案3、18-40、59、60、68、79-98、139、140或146-177中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[NO]紧接在位置570之后存在。

[0599] 179.如实施方案3、18-40、59、60、68、79-98、139、140或146-178中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[NO]置换位置571-573(例如,T571、N572和N573)。

[0600] 180.如实施方案3、18-40、59、60、68、79-98、139、140或146-179中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[NO]紧接在位置570之后存在,且其中[NO]置换位置571-573(例如,T571、N572和N573)。

[0601] 181.如实施方案3、18-40、59、60、68、79-98、139、140或146-181中任一项的AAV衣壳变体,其中[NO]对应于SEQ ID NO:982的位置571-573(例如,T571、N572和N573)。

[0602] 182.如实施方案3、18-40、59、60、68、79-98、139、140或146-181中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[NO]-[N1]-[N2]-[N3]-[N4]紧接在位置570之后存在。

[0603] 183.如实施方案3、18-40、59、60、68、79-98、139、140或146-182中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[NO]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。

[0604] 184.如实施方案3、18-40、59、60、68、79-98、139、140或146-183中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[NO]-[N1]-[N2]-[N3]-[N4]紧接在位置570之后存在,且其中[NO]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。

[0605] 185.如实施方案3、18-40、59、60、68、79-98、139、140或146-184中任一项的AAV衣壳变体,其中[NO]-[N1]-[N2]-[N3]-[N4]对应于SEQ ID NO:982的位置571-586(例如,T571、N572、N573、Q574、S575、S576、Y577、P578、A579、E580、V581、V582、Q583、K584、T585、A586)。

[0606] 186.如实施方案3、31-40、68、94-98、139、140或146-185中任一项的AAV衣壳变体,其中[N4]紧接在根据SEQ ID NO:982编号的位置584之后存在。

[0607] 187.如实施方案3、31-40、68、94-98、139、140或146-186中任一项的AAV衣壳变体,其中[N4]置换位置578和579,这些位置根据SEQ ID NO:138编号;或位置585和586,这些位置根据SEQ ID NO:982编号。

[0608] 188.如实施方案3、31-40、68、94-98、139、140或146-187中任一项的AAV衣壳变体,其中[N4]紧接在位置584之后存在且置换位置585和586,这些位置根据SEQ ID NO:982编号。

[0609] 189.如实施方案3、31-40、68、94-98、139、140或146-188中任一项的AAV衣壳变体,其中:

[0610] (i)[NO]的X_A存在于位置571处,[NO]的X_B存在于位置572处,且[NO]的X_C存在于位置573处,这些位置根据SEQ ID NO:982编号;

[0611] (ii)[N1]的X_D存在于位置574处,[N1]的X_E存在于位置575处,且[N1]的X_F存在于位置576处,这些位置根据SEQ ID NO:982编号;

[0612] (iii)[N2]的X1存在于位置577处,[N2]的X2存在于位置578处,[N2]的X3存在于位置579处,[N2]的X4存在于位置580处,且[N2]的X5存在于位置581处,这些位置根据SEQ ID NO:982编号;

[0613] (iv)[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号;和/或

[0614] (v)[N4]的X_G存在于位置585处且[N4]的X_H存在于位置586处,这些位置根据SEQ ID NO:982编号。

[0615] 190.如实施方案3、31-40、68、94-98、139、140或146-189中任一项的AAV衣壳变体,其中:

[0616] (i)[NO]存在于位置571-573处,这些位置根据SEQ ID NO:982编号;

[0617] (ii)[N1]存在于位置574-576处,这些位置根据SEQ ID NO:982编号;

[0618] (iii)[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号;

[0619] (iv)[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号;

[0620] (v)[N4]存在于位置585-586处,这些位置根据SEQ ID NO:982编号;

[0621] (vi)[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号;和/或

[0622] (vii)[NO]-[N1]-[N2]-[N3]-[N4]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0623] 191.如实施方案1-40、59、60、67-98、139、140或146-190中任一项的AAV衣壳变体,其中[N3]紧接在[N2]之后存在。

[0624] 192.如实施方案3、31-40、68、94-98、139、140或146-191中任一项的AAV衣壳变体,其中[N4]紧接在[N3]之后存在。

[0625] 193.如实施方案1-40、59、60、67-98、139、140或146-192中任一项的AAV衣壳变体,其自N端至C端包含[N2]-[N3]。

[0626] 194.如实施方案1-40、59、60、67-98、139、140或146-193中任一项的AAV衣壳变体,其自N端至C端包含[N1]-[N2]-[N3]。

[0627] 195.如实施方案3、18、19、28-40、59、60、68、89-98、139、140或146-194中任一项的AAV衣壳变体,其自N端至C端包含[NO]-[N1]-[N2]-[N3]。

[0628] 196.如实施方案3、18-40、59、60、68、79-98、139、140或146-195中任一项的AAV衣壳变体,其自N端至C端包含[N1]-[N2]-[N3]-[N4]。

[0629] 197.如实施方案3、18、19、28-40、59、60、68、89-98、139、140或146-196中任一项的AAV衣壳变体,其自N端至C端包含[NO]-[N1]-[N2]-[N3]-[N4]。

[0630] 198.如实施方案41-66或125-145中任一项的AAV衣壳变体,其中[B]-[C]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0631] 199.如实施方案42、52-66、126、134-145或198中任一项的AAV衣壳变体,其中[A]和/或[D]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0632] 200.如实施方案42、52-66、126、134-145、198或199中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0633] 201.如实施方案41-66、125-145或198-200中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]紧接在位置573之后存在。

[0634] 202.如实施方案41-66、125-145或198-201中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]置换位置574-576(例如,Q574、S575和S576)。

[0635] 203.如实施方案41-66、125-145或198-202中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]紧接在位置573之后存在,且其中[B]置换位置574-576(例如,Q574、S575和S576)。

[0636] 204.如实施方案41-66、125-145或198-203中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[B]紧接在位置573之后存在。

[0637] 205.如实施方案41-66、125-145或198-204中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[B]置换位置574-576(例如,Q574、S575和S576)。

[0638] 206.如实施方案41-66、125-145或198-205中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[B]紧接在位置573之后存在,且其中[B]置换位置574-576(例如,Q574、S575和S576)。

[0639] 207.如实施方案41-66、125-145或198-206中任一项的AAV衣壳变体,其中[B]对应于SEQ ID NO:982的位置574-576(例如,Q574、S575和S576)。

[0640] 208.如实施方案41-66、125-145或198-207中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[B]-[C]紧接在位置573之后存在。

[0641] 209.如实施方案41-66、125-145或198-208中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。

[0642] 210.如实施方案41-66、125-145或198-209中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]-[C]紧接在位置573之后存在,且其中[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。

[0643] 211.如实施方案41-66、125-145或198-210中任一项的AAV衣壳变体,其中[B]-[C]对应于SEQ ID NO:982的位置574-584(例如,Q574、S575、S576、Y577、P578、A579、E580、V581、V582、Q583、K584)。

[0644] 212.如实施方案41-66、125-145或198-211中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[B]-[C]-[D]紧接在位置573之后存在。

[0645] 213.如实施方案41-66、125-145或198-212中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]-[C]-[D]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579)。

[0646] 214.如实施方案41-66、125-145或198-213中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[B]-[C]-[D]紧接在位置573之后存在,且其中[B]-[C]-[D]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579)。

[0647] 215.如实施方案41-66、125-145或198-214中任一项的AAV衣壳变体,其中[B]-[C]-[D]对应于SEQ ID NO:982的位置574-586(例如,Q574、S575、S576、Y577、P578、A579、E580、V581、V582、Q583、K584、T585、A586)。

[0648] 216.如实施方案41-66、125-145或198-215中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[C]紧接在位置576之后存在。

[0649] 217.如实施方案41-66、125-145或198-216中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[C]置换位置577(例如,T577)。

[0650] 218.如实施方案41-66、125-145或198-217中任一项的AAV衣壳变体,其中相对于

根据SEQ ID NO:138的氨基酸序列编号的参考序列,[C]紧接在位置576之后存在,且其中[C]置换577(例如,T577)。

[0651] 219.如实施方案41-66、125-145或198-218中任一项的AAV衣壳变体,其中[C]对应于SEQ ID NO:982的位置577-584(例如,Y577、P578、A579、E580、V581、V582、Q583、K584)。

[0652] 220.如实施方案42、52-66、126、134-145或198-219中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[A]紧接在位置570之后存在。

[0653] 221.如实施方案42、52-66、126、134-145或198-220中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[A]置换位置571-573(例如,T571、N572和N573)。

[0654] 222.如实施方案42、52-66、126、134-145或198-221中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[A]紧接在位置570之后存在,且其中[A]置换位置571-573(例如,T571、N572和N573)。

[0655] 223.如实施方案42、52-66、126、134-145或198-222中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[A]紧接在位置570之后存在。

[0656] 224.如实施方案42、52-66、126、134-145或198-223中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[A]置换位置571-573(例如,T571、N572和N573)。

[0657] 225.如实施方案42、52-66、126、134-145或198-224中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:982的氨基酸序列编号的参考序列,[A]紧接在位置570之后存在,且其中[A]置换位置571-573(例如,T571、N572和N573)。

[0658] 226.如实施方案42、52-66、126、134-145或198-225中任一项的AAV衣壳变体,其中[A]对应于SEQ ID NO:982的位置571-573(例如,T571、N572和N573)。

[0659] 227.如实施方案42、52-66、126、134-145或198-226中任一项的AAV衣壳变体,其中[D]紧接在根据SEQ ID NO:982编号的位置584之后存在。

[0660] 228.如实施方案42、52-66、126、134-145或198-227中任一项的AAV衣壳变体,其中[D]置换位置578及579,这些位置根据SEQ ID NO:138编号;或位置585及586,这些位置根据SEQ ID NO:982编号。

[0661] 229.如实施方案42、52-66、126、134-145或198-228中任一项的AAV衣壳变体,其中[D]紧接在位置584之后存在且置换位置585及586,这些位置根据SEQ ID NO:982编号。

[0662] 230.如实施方案42、52-66、126、134-145或198-229中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,[A]-[B]-[C]-[D]紧接在位置570之后存在。

[0663] 231.如实施方案42、52-66、126、134-145或198-230中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[A]-[B]-[C]-[D]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。

[0664] 232.如实施方案42、52-66、126、134-145或198-231中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[A]-[B]-[C]-[D]紧接在位置570之后存在,且其中[A]-[B]-[C]-[D]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。

[0665] 233.如实施方案42、52-66、126、134-145或198-232中任一项的AAV衣壳变体,其中[A]-[B]-[C]-[D]对应于SEQ ID NO:982的位置571-586(例如,T571、N572、N573、Q574、S575、S576、Y577、P578、A579、E580、V581、V582、Q583、K584、T585、A586)。

[0666] 234.如实施方案42、52-66、126、134-145或198-233中任一项的AAV衣壳变体,其中:

[0667] (i)[A]的X_A存在于位置571处,[A]的X_B存在于位置572处,且[A]的X_C存在于位置573处,这些位置根据SEQ ID NO:982编号;

[0668] (ii)[B]的X1存在于位置574处,[B]的X2存在于位置575处,且[B]的X3存在于位置576处,这些位置根据SEQ ID NO:982编号;

[0669] (iii)[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号;和/或

[0670] (iv)[D]的X4存在于位置585处且[D]的位置X5存在于位置586处,这些位置根据SEQ ID NO:982编号。

[0671] 235.如实施方案42、52-66、126、134-145或198-234中任一项的AAV衣壳变体,其中:

[0672] (i)[A]存在于位置571-573处,这些位置根据SEQ ID NO:982编号;

[0673] (ii)[B]存在于位置574-576处,这些位置根据SEQ ID NO:982编号;

[0674] (iii)[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号;

[0675] (iv)[D]存在于位置585-586处,这些位置根据SEQ ID NO:982编号;和/或

[0676] (v)[A]-[B]-[C]-[D]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[0677] 236.如实施方案41-66、125-145或198-235中任一项的AAV衣壳变体,其中[C]紧接在[B]之后存在。

[0678] 237.如实施方案42、61-66、126、141-145或198-236中任一项的AAV衣壳变体,其中[D]紧接在[C]之后存在。

[0679] 238.如实施方案41-66、125-145或198-237中任一项的AAV衣壳变体,其自N端至C端包含[B]-[C]。

[0680] 239.如实施方案42、52-66、126、134-145或198-238中任一项的AAV衣壳变体,其自N端至C端包含[A]-[B]-[C]。

[0681] 240.如实施方案42、61-66、126、141-145或198-239中任一项的AAV衣壳变体,其自N端至C端包含[B]-[C]-[D]。

[0682] 241.如实施方案42、61-66、126、141-145或198-240中任一项的AAV衣壳变体,其自N端至C端包含[A]-[B]-[C]-[D]。

[0683] 242.如实施方案99-124中任一项的AAV衣壳变体,其中[K1]-[K2]存在于环VIII中。

[0684] 243.如实施方案100、92-124或242中任一项的AAV衣壳变体,其中[K0]、[K3]和/或[K4]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0685] 244.如实施方案100、92-124、242或243中任一项的AAV衣壳变体,其中[K0]-[K1]-[K2]-[K3]-[K4]存在于环VIII中,任选地其中环VIII包含位置571-592(例如,氨基酸

TNNQSSTTAPATGTYNLQ EIVP(SEQ ID NO:736)),这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[0686] 245.如实施方案99-124或242-244中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]紧接在位置574之后存在。

[0687] 246.如实施方案99-124或242-245中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]置换位置575-577(例如,S575、S576和T577)。

[0688] 247.如实施方案99-124或242-246中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]紧接在位置574之后存在,且其中[K1]置换位置575-577(例如,S575、S576和T577)。

[0689] 248.如实施方案100、116-124或242-247中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]-[K2]-[K3]紧接在位置574之后存在。

[0690] 249.如实施方案100、116-124或242-248中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]-[K2]-[K3]置换位置575-577(例如,S575、S576和T577)。

[0691] 250.如实施方案100、116-124或242-249中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]-[K2]-[K3]紧接在位置574之后存在,且其中[K1]-[K2]-[K3]置换位置575-577(例如,S575、S576和T577)。

[0692] 251.如实施方案100、121-124或242-250中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]-[K2]-[K3]-[K4]紧接在位置574之后存在。

[0693] 252.如实施方案100、121-124或242-251中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]-[K2]-[K3]-[K4]置换位置575-579(例如,S575、S576、T577、T578和A579)。

[0694] 253.如实施方案100、121-124或242-252中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K1]-[K2]-[K3]-[K4]紧接在位置574之后存在,且其中[K1]-[K2]-[K3]-[K4]置换位置575-579(例如,S575、S576、T577、T578和A579)。

[0695] 254.如实施方案100、110-124或242-253中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K0]紧接在位置570之后存在。

[0696] 255.如实施方案100、110-124或242-254中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K0]置换位置571至574(例如,T571、N572、N573和Q574)。

[0697] 256.如实施方案100、110-124或242-255中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K0]紧接在位置570之后存在,且其中[K0]置换位置571至574(例如,T571、N572、N573和Q574)。

[0698] 257.如实施方案100、121-124或242-256中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K0]-[K1]-[K2]-[K3]-[K4]紧接在位置

570之后存在。

[0699] 258.如实施方案100、121-124或242-257中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K0]-[K1]-[K2]-[K3]-[K4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。

[0700] 259.如实施方案100、121-124或242-258中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,[K0]-[K1]-[K2]-[K3]-[K4]紧接在位置570之后存在,且其中[K0]-[K1]-[K2]-[K3]-[K4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579)。

[0701] 260.如实施方案100、116-124或242-259中任一项的AAV衣壳变体,其中[K3]紧接在[K2]之后存在。

[0702] 261.如实施方案100、121-124或242-260中任一项的AAV衣壳变体,其中[K4]紧接在[K3]之后存在。

[0703] 262.如实施方案100、121-124或242-261中任一项的AAV衣壳变体,其自N端至C端包含[K1]-[K2]。

[0704] 263.如实施方案100、116-124或242-262中任一项的AAV衣壳变体,其自N端至C端包含[K1]-[K2]-[K3]。

[0705] 264.如实施方案100、116-124或242-263中任一项的AAV衣壳变体,其自N端至C端包含[K0]-[K1]-[K2]-[K3]。

[0706] 265.如实施方案100、121-124或242-264中任一项的AAV衣壳变体,其自N端至C端包含[K1]-[K2]-[K3]-[K4]。

[0707] 266.如实施方案100、121-124或242-265中任一项的AAV衣壳变体,其自N端至C端包含[K0]-[K1]-[K2]-[K3]-[K4]。

[0708] 267.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含:

[0709] (a)表1、2A、2B、9或15-20中所提供的任何序列的氨基酸序列;

[0710] (b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;

[0711] (c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0712] (d)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0713] 268.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含:

[0714] (a)SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列;

[0715] (b)包含来自SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;

[0716] (c)相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0717] (d)相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0718] 269.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含:

[0719] (a)SEQ ID NO:943或2064-2080中任一个的氨基酸序列;

[0720] (b)包含来自SEQ ID NO:943或2064-2080中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;

[0721] (c)相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0722] (d)相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0723] 270.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含:

[0724] (a)SEQ ID NO:1021、1024、1232、1300、1327、1533、1538、1585、1590或1591中任一个的氨基酸序列;

[0725] (b)包含来自SEQ ID NO:1021、1024、1232、1300、1327、1533、1538、1585、1590或1591中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;

[0726] (c)相对于SEQ ID NO:1021、1024、1232、1300、1327、1533、1538、1585、1590或1591中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0727] (d)相对于SEQ ID NO:1021、1024、1232、1300、1327、1533、1538、1585、1590或1591中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0728] 271.如实施方案267-270中任一项的AAV衣壳变体,其包含来自SEQ ID NO:943或946-966中任一个的至少3、4、5、6或7个连续氨基酸。

[0729] 272.如实施方案267-271的AAV衣壳变体,其中3个连续氨基酸包含YPA。

[0730] 273.如实施方案267-272的AAV衣壳变体,其中4个连续氨基酸包含YPAE(SEQ ID NO:21)。

[0731] 274.如实施方案267-273的AAV衣壳变体,其中5个连续氨基酸包含YPAEV(SEQ ID NO:1)。

[0732] 275.如实施方案267-274的AAV衣壳变体,其中6个连续氨基酸包含YPAEVV(SEQ ID NO:725)。

[0733] 276.如实施方案267-275的AAV衣壳变体,其中7个连续氨基酸包含YPAEVVQ(SEQ ID NO:726)。

[0734] 277.如实施方案267-276的AAV衣壳变体,其中所述氨基酸序列包含YPAEVVQK(SEQ ID NO:943)。

[0735] 278.如实施方案267、268或270中任一项的AAV衣壳变体,其中:

[0736] (i)3个连续氨基酸包含YTP;

[0737] (ii)4个连续氨基酸包含YTPS(SEQ ID NO:26);

[0738] (iii)5个连续氨基酸包含YTPSL(SEQ ID NO:7);

[0739] (iv)6个连续氨基酸包含YTPSLV(SEQ ID NO:727);

[0740] (v)7个连续氨基酸包含YTPSLVQ(SEQ ID NO:728);和/或

[0741] (vi)所述氨基酸序列包含YTPSLVQK(SEQ ID NO:952)。

[0742] 279.如实施方案267、268或270中任一项的AAV衣壳变体,其中:

[0743] (i)3个连续氨基酸包含YPP;

[0744] (ii)4个连续氨基酸包含YPPS(SEQ ID NO:22);

[0745] (iii)5个连续氨基酸包含YPPSL(SEQ ID NO:2);

[0746] (iv)6个连续氨基酸包含YPPSLV(SEQ ID NO:729);

[0747] (v)7个连续氨基酸包含YPPSLVQ(SEQ ID NO:732);和/或

[0748] (vi)所述氨基酸序列包含YPPSLVQK(SEQ ID NO:946)。

[0749] 280.如实施方案267、268或270中任一项的AAV衣壳变体,其中:

[0750] (i)3个连续氨基酸包含YPP;

[0751] (ii)4个连续氨基酸包含YPPS(SEQ ID NO:22);

[0752] (iii)5个连续氨基酸包含YPPSL(SEQ ID NO:2);

[0753] (iv)6个连续氨基酸包含YPPSLE(SEQ ID NO:733);

[0754] (v)7个连续氨基酸包含YPPSLEQ(SEQ ID NO:734);和/或

[0755] (vi)所述氨基酸序列包含YPPSLEQK(SEQ ID NO:953)。

[0756] 281.如实施方案267、268或270中任一项的AAV衣壳变体,其中:

[0757] (i)3个连续氨基酸包含YPP;

[0758] (ii)4个连续氨基酸包含YPPS(SEQ ID NO:22);

[0759] (iii)5个连续氨基酸包含YPPSL(SEQ ID NO:2);

[0760] (iv)6个连续氨基酸包含YPPSLV(SEQ ID NO:729);

[0761] (v)7个连续氨基酸包含YPPSLVK(SEQ ID NO:735);和/或

[0762] (vi)所述氨基酸序列包含YPPSLVKK(SEQ ID NO:954)。

[0763] 282.如实施方案267-277中任一项的AAV衣壳变体,其包含相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0764] 283.如实施方案267、268、270或278-281中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:1021、1024、1232、1300、1327、1533、1538、1585、1590或1591中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0765] 284.如实施方案267、268、270或378-281中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:946、952、953或954中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[0766] 285.如实施方案267-277或282中任一项的AAV衣壳变体,其包含相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0767] 286.如实施方案267、268、270、278-281或283中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:1021、1024、1232、1300、1327、1533、1538、1585、1590或1591中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0768] 287.如实施方案267、268、270、278-281或284中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:946、952、953或954中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0769] 288.如实施方案267-277、282或285中任一项的AAV衣壳变体,其包含由以下编码的氨基酸序列:

[0770] (i)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代)但不超过十个修饰(例如,取代)的核苷酸序列;或

[0771] (ii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。

[0772] 289.如实施方案267-277、282、285或289中任一项的AAV衣壳变体,其中编码所述氨基酸序列的核苷酸序列包含:

[0773] (i)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代)但不超过十个修饰(例如,取代)的核苷酸序列;或

[0774] (ii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。

[0775] 290.如实施方案267的AAV衣壳变体,其包含:

[0776] (a)SEQ ID NO:2024-2063中任一个的氨基酸序列;

[0777] (b)包含来自SEQ ID NO:2024-2063中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;

[0778] (c)相对于SEQ ID NO:2024-2063中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0779] (d)相对于SEQ ID NO:2024-2063中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0780] 291.如实施方案267或290的AAV衣壳变体,其中相对于SEQ ID NO:2024-2063中任一个的氨基酸序列包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列的不同氨基酸存在于以下位置中的一处或多处:

[0781] (i)位置1,其中不同氨基酸为T或L;

[0782] (ii)位置2,其中不同氨基酸为N、L、K、A、T或P;

[0783] (iii)位置3,其中不同氨基酸为N、K、L、A、Y或S;

[0784] (iv)位置4,其中不同氨基酸为Q、L、T、S、F、Y、K或A;

[0785] (v)位置5,其中不同氨基酸为S、H、A、M、Q、T、V或F;

[0786] (vi)位置6,其中不同氨基酸为S、P、V、A、Q、L、T、N或M;

[0787] (vii)位置7,其中不同氨基酸为Y、H、S、V、A、L或T;

[0788] (viii)位置8,其中不同氨基酸为D、P、A、Q、F、L、S、H或M;

[0789] (ix)位置9,其中不同氨基酸为F、A、L、D或Q;

[0790] (x)位置10,其中不同氨基酸为T、E、I或S;

[0791] (xi)位置11,其中不同氨基酸为V、A、N或S;

[0792] (xii)位置12,其中不同氨基酸为V、L或P;

[0793] (xiii)位置13,其中不同氨基酸为Q、E或P;

[0794] (xiv)位置14,其中不同氨基酸为K、N、S或L;

[0795] (xv)位置15,其中不同氨基酸为T、V、M或L;和/或

[0796] (xvi)位置16,其中不同氨基酸为A、G或R。

[0797] 292.如实施方案267的AAV衣壳变体,其包含:

[0798] (a)SEQ ID NO:1632-2023中任一个的氨基酸序列;

[0799] (b)包含来自SEQ ID NO:1632-2023中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;

[0800] (c)相对于SEQ ID NO:1632-2023中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0801] (d)相对于SEQ ID NO:1632-2023中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0802] 293.如实施方案268或292的AAV衣壳变体,其中相对于SEQ ID NO:1632-2023中任一个的氨基酸序列包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列的不同氨基酸存在于以下位置中的一处或多处:

[0803] (i)位置1,其中不同氨基酸为T、G、N、S、E、L、Y、V或I;

[0804] (ii)位置2,其中不同氨基酸为D、N、K、E、V、G、R、L、H、F、P、T、A、S、I或Y;

[0805] (iii)位置3,其中不同氨基酸为Y、N、K、T、W、Q、M、V、C、A、L、F、H、G、R、S或P;

[0806] (iv)位置4,其中不同氨基酸为H、Q、P、E、R、K、A、S、V、L、T、D、I、G、M或N;

[0807] (v)位置5,其中不同氨基酸为R、S、K、N、H、G、W、A、P、V、Q、Y、L或F;

[0808] (vi)位置6,其中不同氨基酸为G、S、F、R、W、H、I、C、M、A、Y、K、N、Q、V、P、E、D、T或L;

[0809] (vii)位置7,其中不同氨基酸为D、Y、S、I、H、F、P、K、R、G、L、Q、A、M、T、N、V、W、C或E;

[0810] (viii)位置8,其中不同氨基酸为P、L、Q、T、W、V、G、K、I、Y、N、H、R、D、S、M、A、F或E;

[0811] (ix)位置9,其中不同氨基酸为A、R、T、Q、S、M、L、E、K、V、G、D、N、H、F、P或I;

[0812] (x)位置10,其中不同氨基酸为K、E、Q、H、V、G、R、S、P、I、N、M、A、L、D或T;

[0813] (xi)位置11,其中不同氨基酸为V、A、E、N、R、L、M、T、Q、S、K、C、G、D、Y、P、H、F或I;

[0814] (xii)位置12,其中不同氨基酸为V、P、L、S、T、N、A、G、K、R、I、H、E、Q或M;

[0815] (xiii)位置13,其中不同氨基酸为Q、K、N、A、H、R、T、V、E、I、P、G、S或L;

[0816] (xiv)位置14,其中不同氨基酸为K、E、I、Y、Q、R、G、D、L、N或S;

[0817] (xv)位置15,其中不同氨基酸为S、T、N、Q、I、P、E、G、K、M或H;和/或

[0818] (xvi)位置16,其中不同氨基酸为A、D、L、Y、Q或T。

[0819] 294.如实施方案267-293中任一项的AAV衣壳变体,其中所述氨基酸序列存在于环VIII中。

[0820] 295.如实施方案267-294中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置570、571、572、573、574、575或576之后存在。

[0821] 296.如实施方案267-295中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列置换位置571、572、573、574、575和/或576中的一个、两个、三个、四个、五个或全部(例如,位置T571、N572、N573、Q574、S575、S576、T577、T578和/或A579)。

[0822] 297.如实施方案267-296中任一项的AAV衣壳变体,其中相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置576之后存在。

[0823] 298.如实施方案267-297中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述AAV衣壳变体包含位置577处非T的氨基酸残基。

[0824] 299.如实施方案267-298中任一项的AAV衣壳变体,相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,所述AAV衣壳变体包含位置577处的氨基酸Y。

[0825] 300.如实施方案267-299中任一项的AAV衣壳变体,其包含根据SEQ ID NO:138编号的取代T577Y。

[0826] 301.如实施方案298或299的AAV衣壳变体,其中所述氨基酸序列为或包含YPAEVVQK(SEQ ID NO:943),且相对于根据SEQ ID NO:138或982的氨基酸序列编号的参考序列,起始于位置577处。

[0827] 302.如实施方案267-277、282、285、288、289或294-301中任一项的AAV衣壳变体,其包含位置577处非T的氨基酸,且进一步包含PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138编号的位置577之后存在。

[0828] 303.如实施方案267-277、282、285、288、289或294-302中任一项的AAV衣壳变体,其包含位置577处的氨基酸Y,且进一步包含PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138编号的位置577之后存在。

[0829] 304.如实施方案267-277、282、285、288、289或294-303中任一项的AAV衣壳变体,其包含位置577处的氨基酸Y,且进一步包含PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:982编号的位置577之后(例如,位置578-584处)存在。

[0830] 305.如实施方案267-277、282、285、288、289或294-304中任一项的AAV衣壳变体,其包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列置换位置577(例如,T577)。

[0831] 306.如实施方案267-277、282、285、288、289或294-305中任一项的AAV衣壳变体,其包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列,其中所述氨基酸序列:

[0832] (i)紧接在位置576之后存在;及

[0833] (ii)置换位置577(例如,T577),

[0834] 其中在(i)和(ii)中相对于SEQ ID NO:138编号。

[0835] 307.如实施方案267-277、282、285、288、289或294-306中任一项的AAV衣壳变体,其进一步包含在相对于SEQ ID NO:138或SEQ ID NO:982编号的位置571处非T的氨基酸(例如,I)。

[0836] 308.如实施方案267-277、272、275、278、289或294-307中任一项的AAV衣壳变体,其进一步包含在相对于SEQ ID NO:138或SEQ ID NO:982编号的位置571处的I。

[0837] 309.如实施方案267-277、282、285、288、289或294-308中任一项的AAV衣壳变体,其进一步包含以下一个、两个或全部:位置574处非Q的氨基酸(例如,A或T),位置575处非S

的氨基酸(例如,G),和/或位置576处非S的氨基酸(例如,A、L、K或R),这些位置相对于SEQ ID NO:138或SEQ ID NO:982编号。

[0838] 310.如实施方案267-277、282、285、288、289或294-309中任一项的AAV衣壳变体,其进一步包含:

[0839] (i)位置574处的A,位置575处的G,和位置576处的A,这些位置相对于SEQ ID NO:138或982编号;或

[0840] (ii)位置574处的T和位置576处的L,这些位置相对于SEQ ID NO:138或982编号。

[0841] 311.如实施方案267-277、282、285、288、289或294-309中任一项的AAV衣壳变体,其进一步包含:

[0842] (i)在相对于SEQ ID NO:138或982编号的位置576处的K;或

[0843] (ii)在相对于SEQ ID NO:138或982编号的位置576处的R。

[0844] 312.如实施方案267-277、282、285、288、289或294-310中任一项的AAV衣壳变体,其包含:

[0845] (i)位置574处的A,位置575处的G,位置576处的A,和位置577处的Y,这些位置相对于SEQ ID NO:138或982编号;和

[0846] (ii)PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138或982编号的位置577之后存在。

[0847] 313.如实施方案267-277、282、285、288、289或294-310中任一项的AAV衣壳变体,其包含:

[0848] (i)位置574处的T,位置576处的L,和位置577处的Y,这些位置相对于SEQ ID NO:138或982编号;和

[0849] (ii)PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138或982编号的位置577之后存在。

[0850] 314.如实施方案267-277、282、285、288、289、294-309或311中任一项的AAV衣壳变体,其包含:

[0851] (i)位置576处的K和位置577处的Y,这些位置相对于SEQ ID NO:138或982编号;和

[0852] (ii)PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138或982编号的位置577之后存在。

[0853] 315.如实施方案267-277、282、285、288、289、295-309或311中任一项的AAV衣壳变体,其包含:

[0854] (i)位置576处的R和位置577处的Y,这些位置相对于SEQ IDNO:138或982编号;和

[0855] (ii)PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138或982编号的位置577之后存在。

[0856] 316.如实施方案267-277、282、285、288、289、294-308中任一项的AAV衣壳变体,其包含:

[0857] (i)位置571处的I和位置577处的Y,这些位置相对于SEQ ID NO:138或982编号;和

[0858] (ii)PAEVVQK(SEQ ID NO:20)的氨基酸序列,所述氨基酸序列紧接在相对于SEQ ID NO:138或982编号的位置577之后存在。

[0859] 317.如实施方案267、268、270、278、283、284、286或287中任一项的AAV衣壳变体,

其包含位置577处的Y和氨基酸序列TPSLVQK(SEQ ID NO:53),所述氨基酸序列紧接在位置577之后存在,这些位置均根据SEQ ID NO:138或982编号。

[0860] 318.如实施方案267、268、270、279、283、284、286或287中任一项的AAV衣壳变体,其包含位置577处的Y和氨基酸序列PPSLVQK(SEQ ID NO:47),所述氨基酸序列紧接在位置577之后存在,这些位置均根据SEQ ID NO:138或982编号。

[0861] 319.如实施方案267、268、270、280、283、284、286或287中任一项的AAV衣壳变体,其包含位置577处的Y和氨基酸序列PPSLEQK(SEQ ID NO:54),所述氨基酸序列紧接在位置577之后存在,这些位置均根据SEQ ID NO:138或982编号。

[0862] 320.如实施方案267、268、270、281、283、284、286或287中任一项的AAV衣壳变体,其包含位置577处的Y和氨基酸序列PPSLVKK(SEQ ID NO:55),所述氨基酸序列紧接在位置577之后存在,这些位置均根据SEQ ID NO:138或982编号。

[0863] 321.如前述实施方案中任一项的AAV衣壳变体,其进一步包含在环I、II、IV和/或VI中的修饰(例如,插入、取代(例如,保守取代)和/或缺失)。

[0864] 322.如前述实施方案中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:138的氨基酸序列,包含至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过30、20或10个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0865] 323.如前述实施方案中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:138的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同氨基酸的氨基酸序列。

[0866] 324.如前述实施方案中任一项的AAV衣壳变体,所述AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的氨基酸序列。

[0867] 325.如前述实施方案中任一项的AAV衣壳变体,其包含SEQ ID NO:138的氨基酸序列。

[0868] 326.如前述实施方案中任一项的AAV衣壳变体,所述AAV衣壳变体包含由SEQ ID NO:137的核苷酸序列编码的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的序列。

[0869] 327.如前述实施方案中任一项的AAV衣壳变体,其中编码所述衣壳变体的核苷酸序列包含SEQ ID NO:137的核苷酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的序列。

[0870] 328.如前述实施方案中任一项的AAV衣壳变体,其包含VP1蛋白、VP2蛋白、VP3蛋白或其组合。

[0871] 329.如实施方案1-98、125-242、267-277、282、285、288、289、294-306或321-328中任一项的AAV衣壳变体,所述AAV衣壳变体包含对应于SEQ ID NO:982的位置137-731(例如,VP2)的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的序列。

[0872] 330.如实施方案1-98、125-242、267-277、282、285、288、289、294-306或321-329中任一项的AAV衣壳变体,所述AAV衣壳变体包含对应于SEQ ID NO:982的位置193-731(例如,VP3)的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的序列。

[0873] 331.如实施方案1-321中任一项的AAV衣壳变体,所述AAV衣壳变体包含对应于SEQ ID NO:138的位置137-724(例如,VP2)的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的序列。

[0874] 332.如实施方案1-321中任一项的AAV衣壳变体,所述AAV衣壳变体包含对应于SEQ ID NO:138的位置193-724(例如,VP3)的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的序列。

[0875] 333.如实施方案267-277、282、285、288、289、294-306或321-328中任一项的AAV衣壳变体,其包括包含来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5或6个连续氨基酸的氨基酸序列,其中:

[0876] (i)3个连续氨基酸包含YPA;

[0877] (ii)4个连续氨基酸包含YPAE(SEQ ID NO:21);

[0878] (iii)5个连续氨基酸包含YPAEV(SEQ ID NO:1);

[0879] (iv)6个连续氨基酸包含YPAEVV(SEQ ID NO:725);

[0880] (v)7个连续氨基酸包含YPAEVVQ(SEQ ID NO:726);

[0881] 其中所述AAV衣壳变体包含SEQ ID NO:739的氨基酸序列,或与SEQ ID NO:739至少90%、95%、96%、97%、98%或99%同一的氨基酸序列。

[0882] 334.如实施方案267-277、282、285、288、289、294-306、321-328或333中任一项的AAV衣壳变体,其包括包含来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5或6个连续氨基酸的氨基酸序列,其中:

[0883] (i)3个连续氨基酸包含YPA;

[0884] (ii)4个连续氨基酸包含YPAE(SEQ ID NO:21);

[0885] (iii)5个连续氨基酸包含YPAEV(SEQ ID NO:1);

[0886] (iv)6个连续氨基酸包含YPAEVV(SEQ ID NO:725);

[0887] (v)7个连续氨基酸包含YPAEVVQ(SEQ ID NO:726);

[0888] 其中所述AAV衣壳变体包含:(a)包含SEQ ID NO:138或SEQ ID NO:982的氨基酸序列的VP1蛋白;(b)包含SEQ ID NO:138的位置137-724或SEQ ID NO:982的位置137-731的氨基酸序列的VP2蛋白;(c)包含SEQ ID NO:138的位置193-724或SEQ ID NO:982的位置193-731的氨基酸序列的VP3蛋白;或(d)与(a)-(c)中的任何氨基酸序列具有至少90%(例如,至少约95、96、97、98或99%)序列同一性的氨基酸序列。

[0889] 335.如实施方案267-277、282、285、288、289、294-306、321-328或334中任一项的AAV衣壳变体,相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,所述AAV衣壳变体包含一个、两个或三个但不超过四个不同氨基酸,其中所述AAV衣壳变体包含:

[0890] (a)包含SEQ ID NO:138或SEQ ID NO:982的氨基酸序列的VP1蛋白;

[0891] (b)包含SEQ ID NO:138的位置137-724或SEQ ID NO:982的位置137-731的氨基酸序列的VP2蛋白;

[0892] (c)包含SEQ ID NO:138的位置193-724或SEQ ID NO:982的位置193-731的氨基酸序列的VP3蛋白;或

[0893] (d)与(a)-(c)中的任何氨基酸序列具有至少90%(例如,至少约95、96、97、98或99%)序列同一性的氨基酸序列。

[0894] 336.如实施方案267-277、282、285、288、289、294-306、321-328或333-335中任一项的AAV衣壳变体,相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,所述AAV衣壳变体包含一个、两个或三个但不超过四个不同氨基酸,其中所述AAV衣壳变体包含SEQ ID NO:739的氨基酸序列;或与SEQ ID NO:739的氨基酸序列具有至少90%、95%、96%、97%、98%或99%序列同一性的氨基酸序列。

[0895] 337.如实施方案267-277、282、285、288、289、294-306、321-328或333-336中任一项的AAV衣壳变体,相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,所述AAV衣壳变体包含一个或两个但不超过三个取代,其中所述AAV衣壳变体包含与SEQ ID NO:982的氨基酸序列至少90%(例如,至少约95、96、97、98或99%)同一的氨基酸序列。

[0896] 338.如实施方案1-98、125-241、267-277、282、285、288、289、294-306、321-328或333-337中任一项的AAV衣壳变体,所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的氨基酸序列。

[0897] 339.如实施方案1-98、125-241、267-277、282、285、288、289、294-306、321-328或333-338中任一项的AAV衣壳变体,所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少95%同一性的氨基酸序列。

[0898] 340.如实施方案1-98、125-241、267-277、282、285、288、289、294-306、321-328或333-339中任一项的AAV衣壳变体,所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少98%同一性的氨基酸序列。

[0899] 341.如实施方案1-98、125-241、267-277、282、285、288、289、294-306、321-328或333-340中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:982的氨基酸序列,包含至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过30、20或10个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0900] 342.如实施方案1-98、125-241、267-277、282、285、288、289、294-306、321-328或333-341中任一项的AAV衣壳变体,其包含相对于SEQ ID NO:982的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同氨基酸的氨基酸序列。

[0901] 343.如前述实施方案1-98、125-241、267-277、282、285、288、289、294-306、321-328或333-342中任一项的AAV衣壳变体,其中编码所述衣壳变体的核苷酸序列包含SEQ ID NO:984的核苷酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的核苷酸序列。

[0902] 344.如前述实施方案中任一项的AAV衣壳变体,其中编码所述衣壳变体的核苷酸序列为密码子优化的。

[0903] 345.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含如实施方案267-277、282、285、288、289、294-306、321-328或333-342中任一项的氨基酸序列,且进一步包含与SEQ ID NO:982至少80%(例如,至少约85、90、95、96、97、98或99%)同一的氨基酸序列。

[0904] 346.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含与SEQ ID NO:739的氨基酸序列具有至少90%、95%、96%、97%、98%或99%序列同一性的氨基酸序列,其中所述AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列。

[0905] 347.如实施方案346的AAV衣壳变体,其中所述AAV衣壳变体包含SEQ ID NO:739的

氨基酸序列。

[0906] 348.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含SEQ ID NO:982的氨基酸序列。

[0907] 349.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含由以下编码的氨基酸序列:SEQ ID NO:984的核苷酸序列,或与其至少95%同一的核苷酸序列。

[0908] 350.如实施方案1-66、139、140、146-241、267、268、270-289或294-349中任一项的AAV衣壳变体,相对于包含SEQ ID NO:138的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性。

[0909] 351.如实施方案1-66、139、140、146-241、267、268、270-289或294-350中任一项的AAV衣壳变体,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性。

[0910] 352.如实施方案1-66、139、140、146-241、267、268、270-289或294-351中任一项的AAV衣壳变体,相对于包含SEQ ID NO:139的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性。

[0911] 353.如实施方案1-66、139、140、146-241、267、268、270-289或294-352中任一项的AAV衣壳变体,其转导脑区,例如颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团和/或小脑,任选地其中如与SEQ ID NO:139的参考序列相比,转导水平为至少1.5、2.2、2.4、2.5、2.6、2.7、3.0、3.2、3.5、3.7、4.0、4.2、4.5、4.7、4.9、5、10、15、20、25、30、35倍大,例如当通过例如免疫组织化学测定或qPCR测定的测定测量时,例如,如实施例2中所述。

[0912] 354.如实施方案1-66、139、140、146-241、267、268、270-289或294-353中任一项的AAV衣壳变体,与SEQ ID NO:138的参考序列相比,所述AAV衣壳变体在脑中富集至少约2、3、4、5、6、7、8、9、10、12、15、17、20、25、30、35、40、45、50、55、60或65倍,例如当通过如实施例1中所述的测定测量时。

[0913] 355.如实施方案1-66、139、140、146-241、267、268、270-289或294-354中任一项的AAV衣壳变体,与SEQ ID NO:138的参考序列相比,所述AAV衣壳变体在脑中富集至少约10、12、15、17、20、25、30、35、40、45、50、55、60、61、62、63、64或65倍,例如当通过如实施例1中所述的测定测量时。

[0914] 356.如实施方案1-66、139、140、146-241、267、268、270-289或294-355中任一项的AAV衣壳变体,其在至少两个至三个物种(例如,非人灵长类动物和啮齿动物(例如,大鼠和/或小鼠))的脑中富集,例如,如与SEQ ID NO:138的参考序列相比。

[0915] 357.如实施方案1-66、139、140、146-241、267、268、270-289或294-356中任一项的AAV衣壳变体,与SEQ ID NO:138的参考序列相比,所述AAV衣壳变体在至少两个至三个物种(例如,非人灵长类动物和啮齿动物(例如,大鼠和/或小鼠))的脑中富集至少约2、3、4、5、6、7、8、9、10、12、15、17、20、25、30、35、40、45倍,例如当通过如实施例1、2、4和5中所述的测定测量时。

[0916] 358.如实施方案356或357的AAV衣壳变体,其中所述至少两个至三个物种为食蟹猕猴、西非绿猴、普通狨、大鼠和/或小鼠(例如,BALB/c小鼠)。

[0917] 359.如实施方案1-66、139、140、146-241、267、268、270-289或294-358中任一项的

AAV衣壳变体,与SEQ ID NO:982的参考序列相比,所述AAV衣壳变体在脑中富集至少约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、100、125、150、175、200或225倍,例如当通过如实施例4中所述的测定测量时。

[0918] 360.如实施方案1-66、139、140、146-241、267、268、270-289或294-359中任一项的AAV衣壳变体,其将提高水平的有效载荷递送至脑区,任选地其中如与SEQ ID NO:139的参考序列相比,有效载荷的水平提高至少20、25、30、35倍,例如当通过例如qRT-PCR或qPCR测定的测定测量时(例如,如实施例2中所述)。

[0919] 361.如实施方案1-66、139、140、146-241、267、268、270-289或294-360中任一项的AAV衣壳变体,其将提高水平的病毒基因组递送至脑区,任选地其中如与SEQ ID NO:139的参考序列相比,病毒基因组的水平提高至少1.5、2.2、2.4、2.5、2.6、2.7、3.0、3.2、3.5、3.7、4.0、4.2、4.5、4.7、4.9或5倍,例如当通过例如qRT-PCR或qPCR测定的测定测量时(例如,如实施例2中所述)。

[0920] 362.如实施方案350-361中任一项的AAV衣壳变体,其中所述脑区为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑或其组合。

[0921] 363.如实施方案1-66、139、140、146-241、267、268、270-289或294-362中任一项的AAV衣壳变体,与SEQ ID NO:139的参考序列相比,所述AAV衣壳变体在脊髓区中富集至少约3、3.5、4.0、4.5、5、5.0、6.0或6.5倍,例如当通过如实施例2中所述的测定测量时。

[0922] 364.如实施方案363中任一项的AAV衣壳变体,其中所述脊髓区为颈区、腰区、胸区或其组合。

[0923] 365.如前述实施方案中任一项的AAV衣壳变体,相对于背根神经节(DRG)中的转导,所述AAV衣壳变体显示脑区中的优先转导。

[0924] 366.如实施方案1-66、139、140、146-241、267、268、270-289或294-365中任一项的AAV衣壳变体,其能够转导神经元细胞。

[0925] 367.如实施方案1-66、139、140、146-241、267、268、270-289或294-366中任一项的AAV衣壳变体,其能够转导非神经元细胞,例如神经胶质细胞(例如,少突胶质细胞)。

[0926] 368.如实施方案1-66、139、140、146-241、267、268、270-289或294-367中任一项的AAV衣壳变体,相对于肝脏中的转导,所述AAV衣壳变体显示脑区中的优先转导。

[0927] 369.如实施方案67-267、269、271-277、282、285、288、289、294-306或321-349中任一项的AAV衣壳变体,相对于包含SEQ ID NO:138的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对心脏细胞或心脏组织(例如,心室或心房)具有增加的趋向性。

[0928] 370.如实施方案67-267、269、271-277、282、285、288、289、294-306、321-349或369中任一项的AAV衣壳变体,相对于包含SEQ ID NO:139的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对心脏细胞或组织(例如,心室或心房)具有增加的趋向性。

[0929] 371.如实施方案67-267、269、271-277、282、285、288、289、294-306、321-349、369或370中任一项的AAV衣壳变体,其将提高水平的有效载荷递送至心脏区域,任选地其中如与SEQ ID NO:139的参考序列相比,有效载荷的水平提高至少1.5、2或2.5倍,例如当通过例如IHC测定或RT-ddPCR测定的测定测量时(例如,如实施例2中所述)。

[0930] 372.如实施方案67-267、269、271-277、282、285、288、289、294-306、321-349或

369-371中任一项的AAV衣壳变体,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对心脏细胞或组织(例如,心室或心房)具有增加的趋向性。

[0931] 373.如实施方案67-267、269、271-277、282、285、288、289、294-306、321-349或369-372中任一项的AAV衣壳变体,与SEQ ID NO:982的参考序列相比,所述AAV衣壳变体在心脏中富集至少约4、5、10、15、20、25、30、35、40、45或50倍,例如当通过如实施例4中所述的测定测量时。

[0932] 374.如实施方案267、290、291、321-328、331、332或344中任一项的AAV衣壳变体,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对肌肉细胞或组织具有增加的趋向性。

[0933] 375.如实施方案267、290、291、321-328、331、332、344或374中任一项的AAV衣壳变体,与SEQ ID NO:982的参考序列相比,所述AAV衣壳变体在肌肉中富集至少约2、3、4、5、10、15、20、25、30或35倍,例如当通过如实施例4中所述的测定测量时。

[0934] 376.如实施方案374或375的AAV衣壳变体,其中肌肉细胞或组织为心肌(例如,心室或心房或两者)、股四头肌或两者。

[0935] 377.如实施方案267、292、293、321-328、331、332或344中任一项的AAV衣壳变体,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,所述AAV衣壳变体对肝细胞或组织具有增加的趋向性。

[0936] 378.如实施方案267、292、293、321-328、331、332、344或377中任一项的AAV衣壳变体,与SEQ ID NO:982的参考序列相比,所述AAV衣壳变体在肝中富集至少约10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、115、120、130、140、150、160、170、180、185或190倍,例如当通过如实施例4中所述的测定测量时。

[0937] 379.一种多核苷酸,其编码如实施方案1-378中任一项的AAV衣壳变体。

[0938] 380.如实施方案379的多核苷酸,其包含:

[0939] (i)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代)但不超过十个修饰(例如,取代)的核苷酸序列;

[0940] (ii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;或

[0941] (iii)SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。

[0942] 381.如实施方案379或380的多核苷酸,所述多核苷酸包含SEQ ID NO:984的核苷酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的核苷酸序列。

[0943] 382.一种编码AAV衣壳变体(例如,AAV5衣壳变体)的多核苷酸,其中所述多核苷酸包含SEQ ID NO:984的核苷酸序列。

[0944] 383.一种编码AAV衣壳变体(例如,AAV5衣壳变体)的多核苷酸,其中所编码的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列。

[0945] 384.一种编码AAV衣壳变体(例如,AAV5衣壳变体)的多核苷酸,其中所编码的AAV衣壳变体包含:

[0946] (a)表1、2A、2B、9或15-20中所提供的任一序列的氨基酸序列;

[0947] (b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或

[0948] (c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0949] (d)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0950] 385.一种编码AAV衣壳变体(例如,AAV5衣壳变体)的多核苷酸,其中所编码的AAV衣壳变体包含:

[0951] (i)YPAEVVQK(SEQ ID NO:943)的氨基酸序列;

[0952] (ii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;

[0953] (iii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的氨基酸序列;或

[0954] (iv)来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5、6或7个连续氨基酸。

[0955] 386.如实施方案384或385中任一项的多核苷酸,其包含:

[0956] (i)SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;

[0957] (ii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;或

[0958] (iii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[0959] 387.如实施方案384-386中任一项的多核苷酸,其中所述AAV衣壳变体包含:

[0960] (i)SEQ ID NO:982的氨基酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的氨基酸序列;

[0961] (ii)相对于SEQ ID NO:982的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0962] (iii)相对于SEQ ID NO:982的氨基酸序列,包含一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过30、20或10个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。

[0963] 388.如实施方案384-387中任一项的多核苷酸,所述多核苷酸包含SEQ ID NO:984的核苷酸序列,或与其具有至少80%(例如,至少约85、90、95、96、97、98或99%)序列同一性的核苷酸序列。

[0964] 389.如实施方案384-388中任一项的多核苷酸,其包含密码子优化的核苷酸序列。

[0965] 390.一种肽,其包含:

[0966] (a)表1、2A、2B、9或15-20中所提供的任一序列的氨基酸序列;

[0967] (b)包含来自表1、2A、2B、9或15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸的氨基酸序列;或

[0968] (c)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;或

[0969] (d)相对于表1、2A、2B、9或15-20中所提供的任一氨基酸序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。

[0970] 391.一种肽,其包含:

[0971] (i)YPAEVVQK(SEQ ID NO:943)的氨基酸序列;

[0972] (ii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列;

[0973] (iii)相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的氨基酸序列;或

[0974] (iv)来自YPAEVVQK(SEQ ID NO:943)的氨基酸序列的至少3、4、5、6或7个连续氨基酸。

[0975] 392.一种肽,其包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列。

[0976] 393.一种肽,由以下编码:

[0977] (i)SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或

[0978] (ii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;

[0979] (iii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[0980] 394.一种肽,其中编码所述肽的核苷酸序列包含:

[0981] (i)SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;

[0982] (ii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列;或

[0983] (iii)相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[0984] 395.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体包含如实施方案390-394中任一项的肽。

[0985] 396.一种AAV衣壳变体(例如,AAV5衣壳变体),所述AAV衣壳变体由如实施方案382-389中任一项的多核苷酸编码。

[0986] 397.一种AAV颗粒,所述AAV颗粒包含如实施方案1-378、395或396中任一项的AAV衣壳变体。

[0987] 398.如实施方案399的AAV颗粒,所述AAV颗粒包含编码有效载荷的核苷酸序列。

[0988] 399.如实施方案398的AAV颗粒,其中所编码的有效载荷包含治疗性蛋白质或其功能变体;抗体或抗体片段;酶;基因编辑系统的组分;RNAi剂(例如,dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA);或其组合。

[0989] 400.如实施方案399的AAV颗粒,其中所述治疗性蛋白质或其功能变体(例如,重组蛋白)与神经或神经退行性病症、肌肉或神经肌肉病症或神经肿瘤性病症相关(例如,在其中异常表达)。

[0990] 401.如实施方案399或400的AAV颗粒,所述治疗性蛋白质或其功能变体选自载脂蛋白E(APOE)(例如,ApoE2、ApoE3和/或ApoE4);人类运动神经元存活蛋白(SMN)1或SMN2;葡糖脑苷脂酶(GBA1);芳族L-氨基酸脱羧酶(AADC);天冬氨酸酰化酶(ASPA);三肽基肽酶I(CLN2);β-半乳糖苷酶(GLB1);N-磺基葡糖胺磺基水解酶(SGSH);N-乙酰基-α-氨基葡糖苷酶(NAGLU);艾杜糖2-硫酸酯酶(IDS);细胞内胆固醇转运蛋白(NPC1);巨轴索神经蛋白(GAN);或其组合。

[0991] 402.如实施方案399的AAV颗粒,其中所述抗体或抗体结合片段结合至:

[0992] (i)CNS相关靶标,例如与神经或神经退行性病症相关的抗原,例如β-淀粉样蛋白、APOE、tau、SOD1、TDP-43、亨廷顿蛋白(HTT)和/或突触核蛋白;

[0993] (ii)肌肉或神经肌肉相关靶标,例如与肌肉或神经肌肉病症相关的抗原;或

[0994] (iii)神经肿瘤学相关靶标,例如与神经肿瘤性病症相关的抗原,例如HER2或EGFR(例如,EGFRvIII)。

[0995] 403.如实施方案399的AAV颗粒,其中所述酶包含大范围核酸酶、锌指核酸酶、TALEN、重组酶、整合酶、碱基编辑器、Cas9或其片段。

[0996] 404.如实施方案399的AAV颗粒,其中基因编辑系统的所述组分包含CRISPR-Cas系统的一种或多种组分。

[0997] 405.如实施方案404的AAV颗粒,其中所述CRISPR-Cas系统的所述一种或多种组分包含Cas9,例如Cas9直系同源物或Cpf1,和单一指导RNA(sgRNA),任选地其中:

[0998] (i)所述sgRNA位于所述Cas9酶的上游(5');或

[0999] (ii)所述sgRNA位于所述Cas9酶的下游(3')。

[1000] 406.如实施方案399的AAV颗粒,其中所述RNAi剂(例如,dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA)调节(例如,抑制)CNS相关基因、mRNA和/或蛋白质的表达。

[1001] 407.如实施方案406的AAV颗粒,其中所述CNS相关基因选自SOD1、MAPT、APOE、HTT、C9ORF72、TDP-43、APP、BACE、SNCA、ATXN1、ATXN3、ATXN7、SCN1A-SCN5A、SCN8A-SCN11A或其组合。

[1002] 408.如实施方案397-407中任一项的AAV颗粒,其包含病毒基因组,所述病毒基因组包含可操作地连接至编码所述有效载荷的核酸序列的启动子。

[1003] 409.如实施方案408的AAV颗粒,其中所述启动子选自人类延伸因子1α-亚单位(EF1α)、巨细胞病毒(CMV)即刻早期增强子和/或启动子、鸡β-肌动蛋白(CBA)及其衍生物CAG、β葡糖醛酸酶(GUSB)或泛素C(UBC)、神经元特异性烯醇化酶(NSE)、血小板源性生长因子(PDGF)、血小板源性生长因子B链(PDGF-β)、细胞间粘附分子2(ICAM-2)、突触蛋白(Syn)、

甲基-CpG结合蛋白2(MeCP2)、Ca2+/钙调蛋白依赖性蛋白激酶II(CaMKII)、代谢型谷氨酸受体2(mGluR2)、神经丝轻链(NFL)或神经丝重链(NFH)、β-球蛋白微基因nβ2、前脑啡肽原(PPE)、脑啡肽(Enk)和兴奋性氨基酸转运蛋白2(EAAT2)、神经胶质纤维酸性蛋白(GFAP)、髓磷脂碱性蛋白(MBP)、心血管启动子(例如,αMHC、cTnT和CMV-MLC2k)、肝启动子(例如,hAAT、TBG)、骨骼肌启动子(例如,结蛋白、MCK、C512)或功能片段(例如,截短)或其功能变体。

[1004] 410.如实施方案408或409的AAV颗粒,其中所述启动子为EF-1a启动子变体,例如截短的EF-1a启动子。

[1005] 411.如实施方案408-410中任一项的AAV颗粒,其中所述启动子包含SEQ ID NO:2100、2101、2103、2104、2108、2109、2111-2120中任一个的核苷酸序列,或以下任一个:表8中所提供的核苷酸序列,相对于SEQ ID NO:2100、2101、2103、2104、2108、2109、2111-2120的核苷酸序列包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列,或以下任一个:表8中所提供的核苷酸序列,或与SEQ ID NO:2100、2101、2103、2104、2108、2109、2111-2120中任一个具有至少80%(例如,85%、90%、95%、96%、97%、98%或99%)序列同一性的核苷酸序列,或表8中所提供的任一核苷酸序列。

[1006] 412.如实施方案408-411中任一项的AAV颗粒,其中所述病毒基因组进一步包含polyA信号序列。

[1007] 413.如实施方案408-412中任一项的AAV颗粒,其中所述病毒基因组进一步包含反向末端重复(ITR)序列。

[1008] 414.如实施方案408-413中任一项的AAV颗粒,其中所述病毒基因组包含相对于编码所述有效载荷的核酸序列位于5'处的ITR序列。

[1009] 415.如实施方案408-414中任一项的AAV颗粒,其中所述病毒基因组包含相对于编码所述有效载荷的核酸序列位于3'处的ITR序列。

[1010] 416.如实施方案408-415中任一项的AAV颗粒,其中所述病毒基因组包含相对于所述有效载荷位于5'处的ITR序列和相对于编码所述有效载荷的核酸序列位于3'处的ITR序列。

[1011] 417.如实施方案408-416中任一项的AAV颗粒,其中所述病毒基因组进一步包含增强子、Kozak序列、内含子区和/或外显子区。

[1012] 418.如实施方案408-417中任一项的AAV颗粒,其中所述病毒基因组进一步包含编码miR结合位点的核苷酸序列,所述miR结合位点例如调节(例如,减少)由所述病毒基因组编码的有效载荷在其中表达相应miRNA的细胞或组织中的表达的miR结合位点。

[1013] 419.如实施方案418的AAV颗粒,其中所编码的miRNA结合位点与在DRG、肝、心脏、造血或其组合的细胞或组织中表达的miRNA互补,例如完全互补或部分互补。

[1014] 420.如实施方案418或419的AAV颗粒,其中所编码的miR结合位点调节(例如,减少)所编码的抗体分子在DRG、肝、心脏、造血谱系或其组合的细胞或组织中的表达。

[1015] 421.如实施方案408-420中任一项的AAV颗粒,其中所述病毒基因组包含所编码的miR结合位点的至少1-5个拷贝,例如至少1、2、3、4或5个拷贝。

[1016] 422.如实施方案408-421中任一项的AAV颗粒,其中所述病毒基因组包含所编码的miR结合位点的至少3个拷贝,任选地其中所有三个拷贝包含相同miR结合位点,或这些拷贝

中的至少一个、两个、三个或全部包含不同miR结合位点。

[1017] 423.如实施方案422的AAV颗粒,其中所编码的miR结合位点的3个拷贝为连续的(例如,未由间隔子分开)。

[1018] 424.如实施方案422的AAV颗粒,其中所编码的miR结合位点的3个拷贝由间隔子分开,任选地其中所述间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1019] 425.如实施方案408-424中任一项的AAV颗粒,其中所述病毒基因组包含所编码的miR结合位点的至少4个拷贝,任选地其中所有四个拷贝包含相同miR结合位点,或这些拷贝中的至少一个、两个、三个或全部包含不同miR结合位点。

[1020] 426.如实施方案425的AAV颗粒,其中所编码的miR结合位点的4个拷贝为连续的(例如,未由间隔子分开)。

[1021] 427.如实施方案425的AAV颗粒,其中所编码的miR结合位点的4个拷贝由间隔子分开,任选地其中所述间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1022] 428.如实施方案408-427中任一项的AAV颗粒,其中所编码的miR结合位点包含miR122结合位点、miR183结合位点、miR-1结合位点、miR-142-3p或其组合,任选地其中:

[1023] (i)所编码的miR122结合位点包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1024] (ii)所编码的miR183结合位点包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1025] (iii)所编码的miR-1结合位点包含SEQ ID NO:4679的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4679,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;和/或

[1026] (iv)所编码的miR-142-3p结合位点包含SEQ ID NO:4675的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4675,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1027] 429.如实施方案408-428中任一项的AAV颗粒,其中所述病毒基因组包含所编码的

miR122结合位点。

[1028] 430.如实施方案408-429中任一项的AAV颗粒,其中所述病毒基因组包含miR122结合位点的至少1-5个拷贝,例如1、2或3个拷贝,任选地其中每个拷贝为连续的(例如,未由间隔子分开),或每个拷贝由间隔子分开,任选地其中所述间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1029] 431.如实施方案429或430的AAV颗粒,其中所编码的miR122结合位点包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1030] 432.如实施方案408-431中任一项的AAV颗粒,其中所述病毒基因组包含:

[1031] (A)(i)第一所编码的miR122结合位点,其包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1032] (ii)第一间隔子,其包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,包含一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的核苷酸序列;和

[1033] (iii)第二所编码的miR122结合位点,其包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;或

[1034] (B)(i)第一所编码的miR122结合位点,其包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1035] (ii)第一间隔子,其包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,包含一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的核苷酸序列;

[1036] (iii)第二所编码的miR122结合位点,其包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1037] (iv)第二间隔子,其包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,包含一个、

两个或三个修饰(例如,取代(例如,保守取代))但不超过四个修饰(例如,取代(例如,保守取代))的核苷酸序列;和

[1038] (v)第三所编码的miR122结合位点,其包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1039] 433.如实施方案408-432中任一项的AAV颗粒,其中所述病毒基因组包含所编码的miR183结合位点。

[1040] 434.如实施方案408-433中任一项的AAV颗粒,其中所述病毒基因组包含miR183结合位点的至少1-5个拷贝,例如1、2或3个拷贝,任选地其中每个拷贝均为连续的(例如,未由间隔子分开),或每个拷贝由间隔子分开,任选地其中所述间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1041] 435.如实施方案433或434的AAV颗粒,其中所编码的miR183结合位点包含SEQ ID NO:4673的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4673,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1042] 436.如实施方案408-435中任一项的AAV颗粒,其中所述病毒基因组包含:

[1043] (A)(i)第一所编码的miR183结合位点,其包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1044] (ii)第一间隔子,其包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列;和

[1045] (iii)第二所编码的miR183结合位点,其包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;或

[1046] (B)(i)第一所编码的miR183结合位点,其包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1047] (ii)第一间隔子,其包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,包含至少一

个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1048] (iii)第二所编码的miR183结合位点,其包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列;

[1049] (iv)第二间隔子,其包含GATAGTTA的核苷酸序列,或相对于GATAGTTA,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列;和

[1050] (v)第三所编码的miR183结合位点,其包含SEQ ID NO:4676的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列;或相对于SEQ ID NO:4676,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1051] 437.如实施方案408-436中任一项的AAV颗粒,其中所述病毒基因组包含所编码的miR122结合位点和miR-1结合位点。

[1052] 438.如实施方案408-437中任一项的AAV颗粒,其中所述病毒基因组为单链的或自身互补的。

[1053] 439.如实施方案408-438中任一项的AAV颗粒,其中所述病毒基因组进一步包含编码Rep蛋白(例如,非结构蛋白)的核苷酸序列,其中所述Rep蛋白包含Rep78蛋白、Rep68、Rep52蛋白和/或Rep40蛋白(例如,Rep78和Rep52蛋白)。

[1054] 440.如实施方案397-439中任一项的AAV颗粒,其中所述AAV颗粒进一步包含编码Rep蛋白(例如,非结构蛋白)的核苷酸序列,其中所述Rep蛋白包含Rep78蛋白、Rep68、Rep52蛋白和/或Rep40蛋白(例如,Rep78和Rep52蛋白)。

[1055] 441.如实施方案439或440的AAV颗粒,其中所述Rep78蛋白、所述Rep68蛋白、所述Rep52蛋白和/或所述Rep40蛋白由至少一个Rep基因编码。

[1056] 442.如实施方案408-441中任一项的AAV颗粒,其中所述病毒基因组进一步包含编码如实施方案1-364、381或382中任一项的AAV衣壳变体的核苷酸序列。

[1057] 443.如实施方案397-442中任一项的AAV颗粒,其中所述AAV颗粒进一步包含编码如实施方案1-364、381或382中任一项的AAV衣壳变体的核苷酸序列。

[1058] 444.如前述实施方案中任一项的AAV衣壳变体、多核苷酸、肽或AAV颗粒,其为分离的,例如重组的。

[1059] 445.一种载体,其包含编码如实施方案1-378、395、396或444中任一项的AAV衣壳变体的多核苷酸、如实施方案379-389或444中任一项的多核苷酸或编码如实施方案390-394或444中任一项的肽的多核苷酸。

[1060] 446.一种细胞,例如宿主细胞,所述细胞包含如实施方案1-378、395、396或444中任一项的AAV衣壳变体、如实施方案379-389或444中任一项的多核苷酸、如实施方案390-394或444中任一项的肽实施方案、如实施方案397-444中任一项的AAV颗粒或如实施方案

445的载体。

[1061] 447.如实施方案446的细胞,其中所述细胞为哺乳动物细胞或昆虫细胞。

[1062] 448.如实施方案446或447的细胞,其中所述细胞为脑区或脊髓区的细胞,任选地为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑、颈部脊髓、胸部脊髓、腰部脊髓或其组合的细胞。

[1063] 449.如实施方案446-448中任一项的细胞,其中所述细胞为神经元、感觉神经元、运动神经元、星形胶质细胞、神经胶质细胞、少突胶质细胞或肌肉细胞(例如,心脏、横膈膜或股四头肌的细胞)。

[1064] 450.一种制备AAV颗粒的方法,其包括

[1065] (i)提供包含病毒基因组的宿主细胞;及

[1066] (ii)在适合将所述病毒基因组封入如实施方案1-378、395、396或444中任一项的AAV衣壳变体或由如实施方案379-389或444中任一项的多核苷酸编码的AAV衣壳变体中的条件下培育所述宿主细胞;

[1067] 由此制备所述AAV颗粒。

[1068] 451.如实施方案450的方法,所述方法进一步包括在步骤(i)之前,将包含所述病毒基因组的第一核酸分子引入所述宿主细胞中。

[1069] 452.如实施方案450或451的方法,其中所述宿主细胞包含编码所述衣壳变体的第二核酸。

[1070] 453.如实施方案452的方法,其中在所述第一核酸分子之前、同时或之后将所述第二核酸分子引入所述宿主细胞中。

[1071] 454.一种药物组合物,其包含如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒,以及药学上可接受的赋形剂。

[1072] 455.一种将有效载荷递送至细胞或组织(例如,CNS细胞、CNS组织、心脏细胞、心脏组织、肌肉细胞、肌肉组织、肝细胞或肝组织)的方法,其包括施用有效量的如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒。

[1073] 456.如实施方案455的方法,其中所述细胞为脑区或脊髓区的细胞,任选地为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑、脊髓颈区、脊髓胸区、脊髓腰区或其组合的细胞。

[1074] 457.如实施方案455的方法,其中所述细胞为心脏(例如,心房或心室)的细胞。

[1075] 458.如实施方案455的方法,其中所述细胞为肌肉细胞(例如,股四头肌的细胞)或肝细胞。

[1076] 459.如实施方案455或456的方法,其中所述细胞为神经元、感觉神经元、运动神经元、星形胶质细胞、神经胶质细胞或少突胶质细胞。

[1077] 460.如实施方案455-459中任一项的方法,其中所述细胞或组织在受试者内。

[1078] 461.如实施方案460的方法,其中所述受试者患有、已被诊断患有遗传病症(例如,单基因病症或多基因病症),或处于患有遗传病症(例如,单基因病症或多基因病症)的风险

中。

[1079] 462.如实施方案460或461的方法,其中所述受试者患有、已被诊断患有神经(例如,神经退行性)病症,或处于患有神经(例如,神经退行性)病症的风险中。

[1080] 463.如实施方案460或461的方法,其中所述受试者患有、已被诊断患有肌肉病症或神经肌肉病症,或处于患有肌肉病症或神经肌肉病症的风险中。

[1081] 464.如实施方案460或461的方法,其中所述受试者患有、已被诊断患有神经肿瘤性病症,或处于患有神经肿瘤性病症的风险中。

[1082] 465.一种治疗患有或被诊断患有遗传病症(例如,单基因病症或多基因病症)的受试者的方法,所述方法包括向所述受试者施用有效量的如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒。

[1083] 466.一种治疗患有或被诊断患有神经病症(例如,神经退行性病症)的受试者的方法,所述方法包括向所述受试者施用有效量的如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒。

[1084] 467.一种治疗患有或被诊断患有肌肉病症或神经肌肉病症的受试者的方法,所述方法包括向所述受试者施用有效量的如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒。

[1085] 468.一种治疗患有或被诊断患有神经肿瘤性病症的受试者的方法,所述方法包括向所述受试者施用有效量的如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒。

[1086] 469.如实施方案465-468中任一项的方法,其中所述遗传病症、神经病症、神经退行性病症、肌肉病症、神经肌肉病症或神经肿瘤性病症为亨廷顿氏病、肌萎缩性侧索硬化(ALS)、戈谢病、路易体痴呆、帕金森氏病、脊髓性肌萎缩、阿尔茨海默氏病、脑白质营养不良(例如,亚历山大病、常染色体显性脑白质营养不良伴自主神经疾病(ADLD)、卡纳万病(Canavan disease)、脑腱黄瘤病(CTX)、异染性脑白质营养不良(MLD)、佩-梅二氏病(Pelizaeus-Merzbacher disease)或雷夫叙姆病(Refsum disease))或癌症(例如,HER2/neu阳性癌症或胶质母细胞瘤)。

[1087] 470.如实施方案465-469中任一项的方法,其中治疗包括预防所述受试者的所述疾病或病症的进展。

[1088] 471.如实施方案460-470中任一项的方法,其中所述受试者为人。

[1089] 472.如实施方案465-471中任一项的方法,其中所述AAV颗粒被静脉内、经由大池内注射(ICM)、脑内、鞘内、脑室内、经由实质内施用或肌肉内施用于所述受试者。

[1090] 473.如实施方案465-472中任一项的方法,其中所述AAV颗粒经由聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合施用于所述受试者。

[1091] 474.如实施方案465-472中任一项的方法,其中所述AAV颗粒静脉内施用于所述受

试者。

[1092] 475.如实施方案465-475中任一项的方法,其中所述AAV颗粒的施用导致基因、mRNA、蛋白质或其组合的存在、水平和/或活性降低。

[1093] 476.如实施方案465-475中任一项的方法,其中所述AAV颗粒的施用导致基因、mRNA、蛋白质或其组合的存在、水平和/或活性增加。

[1094] 477.如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒,用于将有效载荷递送至细胞或组织的方法中。

[1095] 478.如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒,用于治疗遗传病症、神经病症、神经退行性病症、肌肉病症、神经肌肉病症或神经肿瘤性病症的方法中。

[1096] 479.如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒、包含如实施方案376-380或430中任一项的肽的AAV颗粒,用于制造药剂。

[1097] 480.如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒用于制造药剂的用途。

[1098] 481.如实施方案454的药物组合物、如实施方案397-444中任一项的AAV颗粒、包含如实施方案1-378、395、396或444中任一项的衣壳变体的AAV颗粒或包含如实施方案390-394或444中任一项的肽的AAV颗粒用于制造供治疗遗传病症、神经病症、神经退行性病症、肌肉病症、神经肌肉病症或神经肿瘤性病症用的药剂的用途。

[1099] 本公开的一个或多个实施方案的详情阐述于以下随附描述中。本发明的其他特征、目的和优点将由该描述显而易见。在本说明书中,单数形式还包括复数形式,除非上下文另有明确规定。某些术语在定义部分和全文中定义。

附图说明

[1100] 图1A-1D显示在静脉内施用包含TTN-002衣壳变体(顶部图)或AAV9对照衣壳(底部图)及编码融合至由异源组成型启动子驱动的HA标签的有效载荷的自身互补基因组的AAV颗粒后28天,从NHP(食蟹猕猴)中分离的各种CNS和外周组织的免疫组织化学影像。图1A自左至右显示小脑(浦肯野层)、脊髓(颈部)、皮层(颞叶)和脑干。图1B自左至右显示苍白球、海马体、丘脑、壳核和齿状核。图1C自左至右显示全脑(H层级)、全脑(K层级)和小脑。图1D自左至右显示脊髓(胸部)、DRG(胸部)、肝和心脏。

具体实施方式

[1101] 本文尤其描述包含AAV衣壳变体(例如,本文所述的AAV衣壳变体)的组合物,及其制备和使用方法。通常,AAV衣壳变体对细胞或组织具有增强的趋向性,例如,用于将有效载荷递送至该细胞或组织,例如CNS组织、CNS细胞、心脏细胞、心脏组织、肌肉细胞、肌肉组织、

肝细胞或肝组织。

[1102] 如本文中以下实施例所证实,本文所述的某些AAV衣壳变体显示优于野生型AAV5和/或野生型AAV9的多种优势,包括(i)在静脉内施用后增加的穿透血脑障壁,(ii)更广泛分布于多个脑区,例如额叶皮层、感觉皮层、运动皮层、壳核、丘脑、小脑皮层、齿状核、尾状核和/或海马体,(iii)多个脑区中升高的有效载荷表达,(iv)更广泛分布于一个或多个外周组织(例如,心脏、肌肉和/或肝)中,和/或(v)一个或多个外周组织中升高的有效载荷表达。不希望受理论束缚,据信这些优势可能部分归因于AAV衣壳变体通过脑血管系统的传播。在一些实施方案中,本文所述的AAV衣壳增强有效载荷向脑中包括例如额叶皮层、感觉皮层、运动皮层、壳核、丘脑、小脑皮层、齿状核、尾状核和/或海马体在内的多个区域的递送。在一些实施方案中,本文所述的AAV衣壳增强有效载荷向心脏、肌肉和/或肝的递送。

[1103] 在一些实施方案中,本文所公开的AAV衣壳变体包含AAV5的环VIII中的修饰,例如在相对于SEQ ID NO:138编号的571-579之间的位置处,例如在位置577处。不希望受理论束缚,据信在一些实施方案中,AAV5衣壳的上述区域(例如,571-579之间的位置,例如在位置577处)突出超过3倍对称轴,例如为AAV5衣壳中的表面暴露位置,例如,如Govindasamy等人“Structural Insights into Aden o-Associated Virus Serotype 5,”Journal of Virology,2013,87(20):11187-11199(其内容以引用方式整体并入本文中)所述。在一些实施方案中,环(例如,环VIII)在本文中可与术语可变区(例如,可变区VII I)或VR(例如,VR-VIII)互换使用。在一些实施方案中,环VIII(例如,VR-VIII)包含根据SEQ ID NO:138编号的位置571-592(例如,氨基酸TNNQSSTTAPATGTYNLQEIVP(SEQ ID NO:736))。在一些实施方案中,环VIII(例如,VR-VIII)包含根据SEQ ID NO:982编号的位置571-599(例如,氨基酸TNNQSSYPAEVVQKTAPATGT YNLQEIVP(SEQ ID NO:756))。在一些实施方案中,环VIII或可变VIII(VR-VIII)如Govindasamy等人(同上)(其内容以引用方式整体并入本文中)所述。

[1104] 已使用数种方法来产生对细胞或组织(例如,CNS细胞或组织)具有增强的趋向性的AAV衣壳。一种方法使用培养细胞(Grimm等人In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses.J.Virol.2008年6月82(12):5887-5911)或原位动物组织(Lisowski等人Selection and evaluation of clinically relevant AAV variants in a xenograft liver model.Nature 2014 506:382-386)与腺病毒的共感染,以便触发传染性AAV DNA的指数复制。另一方法涉及使用细胞特异性CRE转基因小鼠(Deverman等人Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain.Nat Biotechnol.2016年2月34(2)204-209;其以引用方式并入本文中),从而允许星形胶质细胞中的特异性病毒DNA重组,随后回收CRE重组的衣壳变体。其他方法应用高通量DNA合成、多路复用、测序技术和机器学习来评估病毒DNA在不同组织中的测序读数,以设计变异型衣壳。这些方法不同于本文所公开的方法。

[1105] 本领域中已知的衣壳生成方法存在一些限制。例如,Deverman等人(2016)所使用的转基因CRE系统在其他动物物种中的能力有限,且小鼠组织中通过定向演化选择的AAV变体在大型动物中未显示出相似特性。先前所述的转导特异性方法不适用于大型动物研究,因为:1)许多相关组织(例如,CNS)不易经受腺病毒共感染,2)特异性腺病毒趋向性自身会使文库分布产生偏向,且3)大型动物通常不适合进行基因转殖或基因工程改造以在规定细

胞类型中表达CRE重组酶。

[1106] 为了解决这些限制,已开发一种广泛适用的功能性AAV衣壳文库筛选平台以在非转基因动物中进行细胞类型特异性生物淘选,且在随附实例中进行描述。在TRACER(通过RNA的细胞类型特异性表达对AAV进行趋向性复位向)平台系统中,将衣壳基因置于细胞类型特异性启动子的控制下,以在不存在辅助病毒共感染的情况下驱动衣壳mRNA表达。不希望受理论束缚,据信此RNA驱动的筛选增加有利于转导特定细胞类型的衣壳变体的选择压力。TRACER平台允许生成AAV衣壳文库,由此实现转导细胞中表达的衣壳mRNA的特异性回收和亚克隆,而无需转基因动物或辅助病毒共感染。不希望受理论束缚,据信由于mRNA转录为完全转导的标志,本文所公开的方法允许鉴定完全传染性AAV衣壳突变体,且除了其较高严格性以外,此方法还允许使用经设计以在任何细胞特异性启动子(诸如但不限于突触蛋白-1启动子(神经元)、GFAP启动子(星形胶质细胞)、TBG启动子(肝)、CAMK启动子(骨骼肌)、MYH6启动子(心肌细胞))控制下表达CAP mRNA的文库来鉴定对特定细胞类型具有高趋向性的衣壳。本文描述使用TRACER方法生成的AAV衣壳变体,这些AAV衣壳变体在例如CNS细胞、CNS组织、肌肉细胞或肌肉组织中证实增强趋向性。

[1107] 本公开的AAV颗粒和有效载荷可递送至一个或多个靶细胞、组织、器官或生物体。在一些实施方案中,本公开的AAV颗粒对靶细胞类型、组织或器官展现增强的趋向性。作为非限制性实例,所述AAV颗粒可对中枢或周围神经系统(分别为CNS和PNS)的细胞和组织具有增强的趋向性。在一些实施方案中,本公开的AAV颗粒可另外或替代地对细胞类型、组织或器官具有降低的趋向性。

[1108] 在一些实施方案中,AAV颗粒由于相对简单结构、其能够感染多种细胞(包括静止细胞和分裂细胞)而无需整合到宿主基因组中且无需复制及其相对良性免疫原性型态而用作生物工具。可操纵病毒的基因组以含有最少的用于组装功能性重组病毒或病毒颗粒的组分,所述功能性重组病毒或病毒颗粒装载有或被工程改造以靶向特定组织且表达或递送所需的有效载荷。

[1109] 在一些实施方案中,所述AAV颗粒为天然存在的(例如,野生型)AAV或重组AAV。在一些实施方案中,野生型AAV病毒基因组为线性、单链DNA(ssDNA)分子,其长度大约为5,000个核苷酸(nt)。在一些实施方案中,反向末端重复序列(ITR)在5'和3'端对病毒基因组加帽,从而为病毒基因组提供复制起点。在一些实施方案中,AAV病毒基因组通常包含两个ITR序列。这些ITR具有特征性T形发夹结构,该结构由处于ssDNA的5'和3'端的自身互补区域(野生型AAV中为145nt)定义,自身互补区域形成能量稳定的双链区域。双链发夹结构包含多种功能,包括但不限于通过充当宿主病毒复制细胞的内源性DNA聚合酶复合物的引子而充当DNA复制起点。

[1110] 在一些实施方案中,野生型AAV病毒基因组进一步包含两个开放阅读框的核苷酸序列,一个开放阅读框用于四种非结构Rep蛋白(Rep78、Rep68、Rep52、Rep40,由Rep基因编码),且一个开放阅读框用于三种衣壳或结构蛋白(VP1、VP2、VP3,由衣壳基因或Cap基因编码)。Rep蛋白用于复制和包装,而衣壳蛋白被组装以产生AAV的蛋白质壳,或AAV衣壳多肽,例如AAV衣壳变体。替代剪接及交替起始密码子和启动子导致由单一开放阅读框生成四种不同Rep蛋白,且由单一开放阅读框生成三种衣壳蛋白。尽管其因AAV血清型而异,但作为非限制性实例,对于AAV5(SEQ ID NO:138和137),VP1是指氨基酸1-724,VP2是指氨基酸137-

724,且VP3是指氨基酸193-724。在一些实施方案中,对于SEQ ID NO:982的氨基酸序列,VP1包含氨基酸1-731,VP2包含氨基酸137-731,且VP3包含氨基酸193-731。换句话说,VP1为全长衣壳序列,而VP2和VP3为整体的较短组分。因此,VP3区域中的序列变化也为VP1和VP2的变化,然而,如与亲本序列相比的百分比差异对于VP3来说将是最大的,因为其为三者中的最短序列。尽管此处关于氨基酸序列进行描述,但编码这些蛋白质的核酸序列可类似地加以描述。三种衣壳蛋白一起组装以产生AAV衣壳蛋白。虽然不希望受理论束缚,但AAV衣壳蛋白通常包含摩尔比为1:1:10的VP1:VP2:VP3。

[1111] 本公开的AAV颗粒可重组产生,且可基于腺相关病毒(AAV)参考序列。除了单链AAV病毒基因组(例如,ssAAV)以外,本公开还提供自身互补的AAV(scAAVs)病毒基因组。scAAV病毒基因组含有DNA链,所述DNA链一起退火以形成双链DNA。通过跳过第二链合成,scAAV允许在转导细胞中的快速表达。在一些实施方案中,本公开的AAV颗粒为scAAV。在一些实施方案中,本公开的AAV颗粒为ssAAV。

[1112] 本领域中公开了用于产生和/或修饰AAV颗粒的方法,诸如假型AAV颗粒(PCT专利公布号WO200028004;WO200123001;WO2004112727;WO2005005610;和WO2005072364,其各自的内容以引用方式整体并入本文中)。

[1113] 如本文所述,包含AAV衣壳变体和病毒基因组的本公开的AAV颗粒对细胞类型或组织(例如,CNS细胞类型、区域或组织)具有增强的趋向性。

[1114] 肽

[1115] 本文公开了肽及包含AAV衣壳变体和肽的相关AAV颗粒,其用于增强或改善靶组织(例如,CNS或PNS的细胞)的转导。在一些实施方案中,所述肽为分离的,例如重组肽。在一些实施方案中,编码所述肽的核酸为分离的,例如重组核酸。

[1116] 在一些实施方案中,所述肽可增加AAV颗粒向CNS的细胞、区域或组织的分布。CNS的细胞可为但不限于神经元(例如,兴奋性、抑制性、运动、感觉、自主、交感、副交感、浦肯野、Betz等)、神经胶质细胞(例如,小神经胶质细胞、星形胶质细胞、少突胶质细胞)和/或脑的支持细胞,诸如免疫细胞(例如,T细胞)。CNS的组织可为但不限于皮层(例如,额叶、顶叶、枕叶、颞叶)、丘脑、下丘脑、纹状体、壳核、尾状核、海马体、内嗅皮层、基底神经节或小脑深部核团。在一些实施方案中,CNS的组织为颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑、颈部脊髓、胸部脊髓或腰部脊髓。

[1117] 在一些实施方案中,所述肽可增加AAV颗粒向PNS的细胞、区域或组织的分布。PNS的细胞或组织可为但不限于背根神经节(DRG)。

[1118] 在一些实施方案中,所述肽可在静脉内施用后增加AAV颗粒向CNS(例如,皮层)的分布。在一些实施方案中,所述肽可在聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合之后增加AAV颗粒向CNS(例如,皮层)的分布。

[1119] 在一些实施方案中,所述肽可在静脉内施用后增加AAV颗粒向PNS(例如,DRG)的分布。在一些实施方案中,所述肽可在聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合之后增加AAV颗粒向PNS(例如,DRG)的分布。

[1120] 在一些实施方案中,所述肽可增加AAV颗粒向心脏(例如,心房或心室)的细胞、区

域或组织的分布。在一些实施方案中,所述肽可在静脉内施用后增加AAV颗粒向心脏细胞、区域或组织的分布。

[1121] 在一些实施方案中,所述肽可增加AAV颗粒向肌肉的细胞、区域或组织的分布。在一些实施方案中,肌肉为心肌(例如,心房或心室)或股四头肌。在一些实施方案中,所述肽可在静脉内施用后增加AAV颗粒向肌肉细胞、区域或组织的分布。

[1122] 在一些实施方案中,所述肽可增加AAV颗粒向肝的细胞、区域或组织的分布。在一些实施方案中,所述肽可在静脉内施用后增加AAV颗粒向肝细胞、区域或组织的分布。

[1123] 肽的长度可不同。在一些实施方案中,所述肽为约3至约20个氨基酸的长度。作为非限制性实例,所述肽可为3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或3-5、3-8、3-10、3-12、3-15、3-18、3-20、5-10、5-15、5-20、10-12、10-15、10-20、12-20或15-20个氨基酸的长度。在一些实施方案中,肽包含约6至12个氨基酸的长度,例如约9个氨基酸的长度。在一些实施方案中,肽包含约7至11个氨基酸的长度,例如约8个氨基酸的长度。在一些实施方案中,肽包含约5至10个氨基酸的长度,例如约7个氨基酸的长度。在一些实施方案中,肽包含约4至9个氨基酸的长度,例如约6个氨基酸的长度。

[1124] 在一些实施方案中,肽可包含如表1中所陈述的序列(例如,包含SEQ ID NO:1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590-1593、1598-1624或2064-2079中任一个的氨基酸序列)。在一些实施方案中,所述肽可包含如表2A或2B中所陈述的序列。在一些实施方案中,所述肽可包含表9或15-20中所陈述的序列。在一些实施方案中,所述肽为分离的,例如重组的。

[1125] 表1.例示性肽序列

[1126]

肽序列	SEQ ID NO:	肽序列	SEQ ID NO:	肽序列	SEQ ID NO:
TNNQSSYPAEVVQKTA	1533	TNNQSSYPPSLEQKTA	1590	TNNQSSYPAEVVQKNA	1619
TNNAGAYPAEVVQKTA	1021	TNNQSSYPPSLVKKTA	1591	TNNQRKYPAEVVQKTA	1620
INNQSSYPAEV	1024	TNNLSSYQAEV	1592	TNNQSSNPAEE	1621

VQKTA		VQKTA		VQKTA	
TNNLGSYPAEV VQKTA	1027	TNNQSSYPPSL VQKPA	1593	NNNQSSYPAEV VQKTA	1622
TNNIGSYPAEV VQKTA	1112	TNNHSSYPAAV VQKTA	1598	TNNQSYYPAEE VQKTA	1623
TNNSSSYPAEV VQKPA	1142	TNSQSSNPAEV VQKTA	1599	TKNHSSYPAEV VQKTA	1624
TNNSSSYTAEV VQKTA	1214	NNNQSRYPAE VVQKTA	1601	TNSSLSYQAEV VQKTA	2064
TNNTSLYPAEV VQKTA	1232	TNNQSSYTAEV VQKNA	1602	TNSSLSYTAEV VQKTA	2065
TNNQSKYPAE VEQKTA	1254	TNNTSLCPAEV VQKTA	1603	TNSSLYYPAEV VQKTA	2066
TNNQSSYPPSL VQKTA	1300	TNNQRSYTAE VVQKTA	1604	TNTSATYPAEV VQKTA	2067
TNNHSSYPAEV VQKTA	1310	TNSTSLYPAEV VQKTA	1605	TNSSLSYPAEV VQKTA	2068
TNNQSRYPAEV VQKTA	1327	TNNQSSYTAEV VKKTA	1606	TNSSLSYPAEE VQKTA	2069
TNNSSSYPAEV VKKTA	1331	TNNHSSYPAEV LQKTA	1607	TNSSLSYPAEV VQKTD	2070
TNNQSRYPAEE VQKTA	1342	TNNQSSYQAEE VQKTA	1608	TNSSLSYPAEV VQNTA	2071
TNNSSSYPAEV VQQTA	1419	TNNKSKYPAE VVQKTA	1610	TNSSLSYPAEV VQKNA	2072
TNNQSKYPAE VVHKTA	1453	TNNQSRYPAEV MQKTA	1611	TNSSLSYPAEV VQKPA	2073
TNNQSKYPAE VVQKTA	1538	TNSTSPYPAEV VQKTA	1612	TNNQTSTEAEV VQKTA	2074
TNNSSSYPAEV VQKTA	1539	TNNQSSYPAW LIQKTA	1613	TNKSATYPAEV VQKTA	2075
TNTASSYPAEV VQKTA	1575	TNNQSSYPAEA TKKTA	1614	TNSSLSYPADV VQKTA	2076
TNNQSTNKAE VVQKTA	1578	TNKQSSYTAEV VQKTA	1615	TNNQSSYPAEV VQKTA	2077
TNNTSLYPAEE VQKTA	1583	TNKIGSYPAEV VQKTA	1616	TNNQSSYPAEV VQKNA	2078
TNTASSYQAEV VQKTA	1584	TNNKSRYPAEV VQKTA	1617	TNNQSSYPSTD VQKTA	2079
TNNQSSYTPSL VQKTA	1585	TNNQSSYPAEV VQKTD	1618	TNNKGLYPAE VVQKTA	2080
TNNQASYPAE VVQKTA	1586	TNKQASYPAE VVQKTA	1587	TNSTHSYPAEV VQKTA	751

[1127]

[1128] 表2A.例示性肽序列

[1129]

肽	SEQ ID NO:	氨基酸序列	SEQ ID NO:	核苷酸序列

[1130]

2	943	YPAEVVQK	944	TATCCGGCGGAGGTGGTGC AGAAG

[1131] 表2B.例示性肽序列

[1132]

氨基酸序列	SEQ ID NO:	氨基酸序列	SEQ ID NO:	氨基酸序列	SEQ ID NO:
YPAEVVQK	943	YPPSLEQK	953	CPAEVVQK	960
YPPSLVQK	946	YPPSLVKK	954	YTAEVVKK	961
NKAEVVQK	947	YPAEVVKK	955	YPAEVLQK	962
YTAEVVQK	948	YPAEVVHK	956	YQAEEVQK	963
YPAEVEQK	949	YPAAVVQK	957	YPAEVVQN	964
YPAEEVQK	950	NPAEVVQK	958	TEAEVVQK	965
YQAEVVQK	951	YPAEVVQQ	959	YPADVVQK	966
YTPSLVQK	952				

[1133] 在一些实施方案中,本文所述的肽包含具有式[N2]-[N3]的氨基酸序列,其中[N2]包含位置X1、X2、X3、X4和X5且[N3]包含VQK、VQN、EQK、VKK、VHK、VQQ或LQK的氨基酸序列。在一些实施方案中,[N2]的位置X1为Y、N、C或T。在一些实施方案中,[N2]的位置X2为P、E、K、T或Q。在一些实施方案中,[N2]的位置X3为A或P。在一些实施方案中,[N2]的位置X4为E、S、D或A。在一些实施方案中,[N2]的位置X5为V、L或E。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N3]的氨基酸序列包含VQK。在一些实施方案中,[N3]的氨基酸序列由VQK组成。

[1134] 在一些实施方案中,本文所述的肽包含具有式[N2]-[N3]的氨基酸序列,其中[N2]包含位置X1、X2、X3、X4和X5且[N3]包含VQK、EQK、VKK、VHK、VQQ或LQK的氨基酸序列。在一些实施方案中,[N3]包含VQK、EQK或VKK的氨基酸序列。在一些实施方案中,[N3]包含氨基酸序列VQK。在一些实施方案中,[N3]由氨基酸序列VQK组成。在一些实施方案中,[N2]的位置X1为Y、N或C。在一些实施方案中,[N2]的位置X1为Y或N。在一些实施方案中,[N2]的位置X2为P、K、T或Q。在一些实施方案中,[N2]的位置X2为P、T或Q。在一些实施方案中,[N2]的位置X3为A或P。在一些实施方案中,[N2]的位置X3为A。在一些实施方案中,[N2]的位置X4为E、S或A。在一些实施方案中,[N2]的位置X5为V、L或E。在一些实施方案中,[N2]的位置X5为V或L。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N2]包含YPA、YPP、NKA、YTA、YQA、YTP、NPA、CPA、THA、PAE、PPS、KAE、TAE、QAE、TPS、PAA、HAS、AEV、PSL、AEE或AAV。在一些实施方案中,[N2]包含YPAE(SEQ ID NO:21)、YPPS(SEQ ID NO:22)、NKAE(SEQ ID NO:23)、YTAE(SEQ ID NO:24)、YQAE(SEQ ID NO:25)、YTPS(SEQ ID NO:26)、YPAA(SEQ ID NO:27)、NPAE(SEQ ID NO:28)、CPAE(SEQ ID NO:29)、THAS(SEQ ID NO:30)、PAEV(SEQ ID NO:17)、PPSL(SEQ ID NO:31)、KAEV(SEQ ID NO:32)、TAEV(SEQ ID NO:16)、PAEE(SEQ ID NO:18)、QAEV(SEQ ID NO:15)、TPSL(SEQ ID NO:33)、PAAV(SEQ ID NO:34)或QAEE(SEQ ID NO:35)。

在一些实施方案中,[N2]为或包含YPAEV(SEQ ID NO:1)、YPPSL(SEQ ID NO:2)、NKAEV(SEQ ID NO:3)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、YQAEV(SEQ ID NO:6)、YTPSL(SEQ ID NO:7)、YPAAV(SEQ ID NO:8)、NPAEV(SEQ ID NO:9)、CPAEV(SEQ ID NO:10)或YQAEE(SEQ ID NO:11)。在一些实施方案中,[N2]包含YPAEV(SEQ ID NO:1)的氨基酸序列。在一些实施方案中,[N2]的氨基酸序列由YPAEV(SEQ ID NO:1)组成。在一些实施方案中,[N2]-[N3]包含AEVVQK(SEQ ID NO:36)、PSLVQK(SEQ ID NO:37)、AEVEQK(SEQ ID NO:38)、AEEVQK(SEQ ID NO:39)、PSLEQK(SEQ ID NO:40)、PSLVKK(SEQ ID NO:41)、AEVVKK(SEQ ID NO:42)、AEVVHK(SEQ ID NO:43)、AAVVQK(SEQ ID NO:44)、AEVVQQ(SEQ ID NO:45)或AEVLQK(SEQ ID NO:46)的氨基酸序列。在一些实施方案中,[N2]-[N3]包含氨基酸序列PAEVVQK(SEQ ID NO:20)、PPSLVQK(SEQ ID NO:47)、KAEVVQK(SEQ ID NO:48)、TAEVVQK(SEQ ID NO:49)、PAEVEQK(SEQ ID NO:50)、PAEEVQK(SEQ ID NO:51)、QAEVVQK(SEQ ID NO:52)、TPSLVQK(SEQ ID NO:53)、PPSLEQK(SEQ ID NO:54)、PPSLVKK(SEQ ID NO:55)、PAEVVKK(SEQ ID NO:56)、PAEVVHK(SEQ ID NO:57)、PAAVVQK(SEQ ID NO:58)、PAEVVQQ(SEQ ID NO:59)、TAEVVKK(SEQ ID NO:60)、PAEVLQK(SEQ ID NO:61)或QAEEVQK(SEQ ID NO:62)。在一些实施方案中,[N2]-[N3]为或包含YPAEVVQK(SEQ ID NO:943)、YPPSLVQK(SEQ ID NO:946)、NKAEVVQK(SEQ ID NO:947)、YTAEVVQK(SEQ ID NO:948)、YPAEVEQK(SEQ ID NO:949)、YPAEEVQK(SEQ ID NO:950)、YQAEVVQK(SEQ ID NO:951)、YTPSLVQK(SEQ ID NO:952)、YPPSLEQK(SEQ ID NO:953)、YPPSLVKK(SEQ ID NO:954)、YPAEVVKK(SEQ ID NO:955)、YPAEVVHK(SEQ ID NO:956)、YPAAVVQK(SEQ ID NO:957)、NPAEVVQK(SEQ ID NO:958)、YPAEVVQQ(SEQ ID NO:959)、CPAEVVQK(SEQ ID NO:960)、YTAEVVKK(SEQ ID NO:961)、YPAEVLQK(SEQ ID NO:962)或YQAEEVQK(SEQ ID NO:963);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N2]-[N3]为YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,[N2]-[N3]包含氨基酸序列YPAEVVQK(SEQ ID NO:943)。

[1135] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的肽进一步包含[N1],其包含位置X_D、X_E和X_F。在一些实施方案中,[N1]的位置X_D为Q、T、S、A、I、L或H。在一些实施方案中,[N1]的位置X_E为S、G、A或R。在一些实施方案中,[N1]的位置X_F为S、K、L、R、A或T。在一些实施方案中,[N1]包含SK、SL、SS、SR、GA、GS、AS、ST、RS、QS、TS、AG、IG、QA、LG、HS、LS或QR。在一些实施方案中,[N1]为或包含QSS、QSK、TSL、SSS、QSR、AGA、IGS、QAS、ASS、LGS、QST、HSS、LSS或QRS。在一些实施方案中,[N1]的氨基酸序列为QSS。在一些实施方案中,[N1]-[N2]包含SSYPA(SEQ ID NO:63)、SKYPA(SEQ ID NO:64)、SLYPA(SEQ ID NO:65)、SRYPA(SEQ ID NO:66)、SSYPP(SEQ ID NO:67)、GAYPA(SEQ ID NO:68)、GSYPA(SEQ ID NO:69)、ASYPA(SEQ ID NO:70)、STNKA(SEQ ID NO:71)、SSYTA(SEQ ID NO:72)、SSYQA(SEQ ID NO:73)、SSYTP(SEQ ID NO:74)、SSNPA(SEQ ID NO:75)、SLCPA(SEQ ID NO:76)、RSYTA(SEQ ID NO:77)或SSTHA(SEQ ID NO:78)。在一些实施方案中,[N1]-[N2]包含SSYPAE(SEQ ID NO:79)、SKYPAE(SEQ ID NO:80)、SLYPAE(SEQ ID NO:81)、SRYPAE(SEQ ID NO:82)、SSYPPS(SEQ ID NO:83)、GAYPAE(SEQ ID NO:84)、GSYPAE(SEQ ID NO:85)、ASYPAE(SEQ ID NO:86)、STNKAE(SEQ ID

NO:87)、SSYTAE(SEQ ID NO:88)、SSYQAE(SEQ ID NO:89)、SSYTPS(SEQ ID NO:90)、SSYPAA(SEQ ID NO:91)、SSNPAE(SEQ ID NO:92)、SLCPAE(SEQ ID NO:93)、RSYTAE(SEQ ID NO:94)、SSTHAS(SEQ ID NO:95)。在一些实施方案中,[N1]-[N2]为或包含QSSYPAEV(SEQ ID NO:96)、QSKYPAEV(SEQ ID NO:97)、TSLYPAEV(SEQ ID NO:98)、SSSYPAEV(SEQ ID NO:99)、QSRYPAEV(SEQ ID NO:100)、QSSYPPSL(SEQ ID NO:101)、AGAYPAEV(SEQ ID NO:102)、IGSYPAEV(SEQ ID NO:103)、QASYPAEV(SEQ ID NO:104)、ASSYPAEV(SEQ ID NO:105)、LGSYPAEV(SEQ ID NO:106)、QSTNKAEV(SEQ ID NO:107)、HSSYPAEV(SEQ ID NO:108)、SSSYTAEV(SEQ ID NO:109)、TSLYPAEE(SEQ ID NO:110)、ASSYQAEV(SEQ ID NO:111)、QSSYTPSL(SEQ ID NO:112)、QSRYPAEE(SEQ ID NO:113)、LSSYQAEV(SEQ ID NO:114)、HSSYPAAV(SEQ ID NO:115)、QSSNPAEV(SEQ ID NO:116)、QSSYTAEV(SEQ ID NO:117)、TSLCPAEV(SEQ ID NO:118)、QRSYTAEV(SEQ ID NO:119)或QSSYQAEE(SEQ ID NO:120);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]的氨基酸序列为QSSYPAEV(SEQ ID NO:96)。在一些实施方案中,[N1]-[N2]-[N3]包含SSYPAEVVQ(SEQ ID NO:121)、SKYPAEVVQ(SEQ ID NO:122)、SLYPAEVVQ(SEQ ID NO:123)、SRYPAEVVQ(SEQ ID NO:124)、SSYPPSLVQ(SEQ ID NO:125)、GAYPAEVVQ(SEQ ID NO:126)、GSYPAEVVQ(SEQ ID NO:127)、ASYPAEVVQ(SEQ ID NO:128)、STNKAEVVQ(SEQ ID NO:129)、SSYTAEVVQ(SEQ ID NO:130)、SKYPAEVEQ(SEQ ID NO:131)、SLYPAEEVQ(SEQ ID NO:132)、SSYQAEVVQ(SEQ ID NO:133)、SSYTPSLVQ(SEQ ID NO:134)、SRYPAEEVQ(SEQ ID NO:135)、SSYPPSLEQ(SEQ ID NO:136)、SSYPPSLVK(SEQ ID NO:140)、SSYPAEVVK(SEQ ID NO:141)、SKYPAEVVH(SEQ ID NO:142)、SSYPAAVVQ(SEQ ID NO:143)、SSNPAEVVQ(SEQ ID NO:144)、SLCPAEVVQ(SEQ ID NO:145)、RSYTAEVVQ(SEQ ID NO:146)、SSYTAEVVK(SEQ ID NO:147)、SSYPAEVLQ(SEQ ID NO:148)或SSYQAEEVQ(SEQ ID NO:149)。在一些实施方案中,[N1]-[N2]-[N3]为或包含QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、QSSYPPSLVQK(SEQ ID NO:155)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)、QSTNKAEVVQK(SEQ ID NO:161)、HSSYPAEVVQK(SEQ ID NO:162)、SSSYTAEVVQK(SEQ ID NO:163)、QSKYPAEVEQK(SEQ ID NO:164)、TSLYPAEEVQK(SEQ ID NO:165)、ASSYQAEVVQK(SEQ ID NO:166)、QSSYTPSLVQK(SEQ ID NO:167)、QSRYPAEEVQK(SEQ ID NO:168)、QSSYPPSLEQK(SEQ ID NO:169)、QSSYPPSLVKK(SEQ ID NO:170)、LSSYQAEVVQK(SEQ ID NO:171)、SSSYPAEVVKK(SEQ ID NO:172)、QSKYPAEVVHK(SEQ ID NO:173)、HSSYPAAVVQK(SEQ ID NO:174)、QSSNPAEVVQK(SEQ ID NO:175)、SSSYPAEVVQQ(SEQ ID NO:176)、QSSYTAEVVQK(SEQ ID NO:177)、TSLCPAEVVQK(SEQ ID NO:178)、QRSYTAEVVQK(SEQ ID NO:179)、QSSYTAEVVKK(SEQ ID NO:180)、HSSYPAEVLQK(SEQ ID NO:181)或QSSYQAEEVQK(SEQ ID NO:182);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超

过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]-[N3]的氨基酸序列为QSSYPAEVVQK(SEQ ID NO:150)。

[1136] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的肽进一步包含[NO],其中[NO]包含位置X_A、X_B和X_C。在一些实施方案中,[NO]的位置X_A为T、I或N。在一些实施方案中,[NO]的位置X_B为N。在一些实施方案中,[NO]的位置X_C为N、T、S或K。在一些实施方案中,[NO]包含TN、IN、NN、NT、NS或NK。在一些实施方案中,[NO]为或包含TNN、TNT、INN、TNS、NNN或TNK。在一些实施方案中,[NO]的氨基酸序列为TNN。在一些实施方案中,[NO]-[N1]为或包含TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNQST(SEQ ID NO:193)、TNNHSS(SEQ ID NO:194)、TNNQSK(SEQ ID NO:184)、TNNLSS(SEQ ID NO:195)、INNQSS(SEQ ID NO:196)、TNSQSS(SEQ ID NO:197)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)、TNNQRS(SEQ ID NO:200)或TNKQAS(SEQ ID NO:201);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[NO]-[N1]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[NO]-[N1]-[N2]-[N3]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVVQK(SEQ ID NO:510)、TNNQSSYPPSLVQK(SEQ ID NO:512)、TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQK(SEQ ID NO:517)、TNTASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNQSTNKAEVVQK(SEQ ID NO:524)、TNNHSSYPAEVVQK(SEQ ID NO:525)、TNNSSSYTAEVVQK(SEQ ID NO:526)、TNNQSKYPAEVEQK(SEQ ID NO:529)、TNNTSLYPAEEVQK(SEQ ID NO:530)、TNTASSYQAEVVQK(SEQ ID NO:531)、TNNQSSYTPSLVQK(SEQ ID NO:533)、TNNQSRYPAEEVQK(SEQ ID NO:534)、TNNQSSYPPSLEQK(SEQ ID NO:535)、TNNQSSYPPSLVKK(SEQ ID NO:536)、TNNLSSYQAEVVQK(SEQ ID NO:539)、TNNSSSYPAEVVKK(SEQ ID NO:540)、TNNQSKYPAEVVHK(SEQ ID NO:542)、INNQSSYPAEVVQK(SEQ ID NO:543)、TNNHSSYPAAVVQK(SEQ ID NO:545)、TNSQSSNPAEVVQK(SEQ ID NO:548)、TNNSSSYPAEVVQQ(SEQ ID NO:551)、NNNQSRYPAEVVQK(SEQ ID NO:552)、TNNQSSYTAEVVQK(SEQ ID NO:553)、TNNTSLCPAEVVQK(SEQ ID NO:554)、TNSTSLYPAEVVQK(SEQ ID NO:556)、TNNQRSYTAEVVQK(SEQ ID NO:557)、TNNQSSYTAEVVKK(SEQ ID NO:558)、TNNHSSYPAEVLQK(SEQ ID NO:560)、TNNQSSYQAEEVQK(SEQ ID NO:562)或TNKQASYPAEVVQK(SEQ ID NO:563);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[NO]-[N1]-[N2]-[N3]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1137] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的肽进一步包含[N4],其包含位置X_G和X_H。在一些实施方案中,[N4]的位置X_G为T、P或N。在一些实施方案中,[N4]的位置X_G为T。在一些实施方案中,[N4]的位置X_H为A。在一些实施方案中,[N4]为或包含TA、PA或NA。在一些实施方案中,[N4]为TA。在一些实施方案中,[N3]-[N4]为或包含VQKTA(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKKTA(SEQ ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SEQ ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571);包含任何上述氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N3]-[N4]为VQKTA(SEQ ID NO:564)。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为或包含TNNQSSYPAEVVQKTA(SE Q ID NO:1533)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEVVQK TA(SEQ ID NO:1539)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)、TNNQSSYPPSLVQKTA(SEQ ID NO:1300)、TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNIGSYPAEVVQKTA(SEQ ID NO:1112)、TNNQASYPAEVVQKTA(SEQ ID NO:1194)、TN TASSYPAEVVQKTA(SEQ ID NO:1575)、TNNLGSYPAEVVQKT A(SEQ ID NO:1027)、TNNQSTNKAEVVQKTA(SEQ ID NO:1578)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNNQSKYPA EVVQKTA(SEQ ID NO:1538)、TNNSSSYTAEVVQKTA(SEQ ID NO:1214)、TNNQSKYPAEVEQKTA(SEQ ID NO:1254)、TNN TSLYPAEEVQKTA(SEQ ID NO:1583)、TNTASSYQAEVVQKTA(SEQ ID NO:1584)、TNNQSSYTPSLVQKTA(SEQ ID NO:1585)、TNNQSRYPAEEVQKTA(SEQ ID NO:1342)、TNNQSSYPPSLEQK TA(SEQ ID NO:1590)、TNNQSSYPPSLVKKTA(SEQ ID NO:1591)、TNNLSSYQAEVVQKTA(SEQ ID NO:1592)、TNNQSSYPPSLVQKPA(SEQ ID NO:1593)、TNNSSSYPAEVVKKTA(SEQ ID NO:1331)、TNNQSKYPAEVVHKTA(SEQ ID NO:1453)、TNNS SSYPAEVVQKPA(SEQ ID NO:1142)、INNQSSYPAEVVQKTA(SEQ ID NO:1024)、TNNHSSYPAAVVQKTA(SEQ ID NO:1598)、TNSQSSNPAEVVQKTA(SEQ ID NO:1599)、TNNSSSYPAEVVQ QTA(SEQ ID NO:1419)、NNNQSRYPAEVVQKTA(SEQ ID NO:1601)、TNNQSSYTAEVVQKNA(SEQ ID NO:1602)、TNNTSLC PAEVVQKTA(SEQ ID NO:1603)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)、TNNQRSYTAEVVQKTA(SEQ ID NO:1604)、TNNQSSYTAEVVKKTA(SEQ ID NO:1606)、TNNHSSYPAEVLQKT A(SEQ ID NO:1607)、TNNQSSYQAEEVQKTA(SEQ ID NO:1608)或TNKQASYPAEVVQKTA(SEQ ID NO:1587);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1138] 在一些实施方案中,本文所述的肽包含式[N2]-[N3],其中[N2]包含位置X1、X2、X3、X4和X5且[N3]包含VQK或VQN的氨基酸序列。在一些实施方案中,[N3]包含氨基酸序列VQK。在一些实施方案中,[N2]的位置X1为Y或T。在一些实施方案中,[N2]的位置X2为Q、T、P或E。在一些实施方案中,[N2]的位置X3为A。在一些实施方案中,[N2]的位置X4为E或D。在一

些实施方案中,[N2]的位置X4为E或D。在一些实施方案中,[N2]的位置X5为V或E。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N2]包含YP、YQ、YT、TE、QA、TA、PA、EA、EV、EE、DV、AE或AD。在一些实施方案中,[N2]包含YPA、YQA、YTA、TEA、QAE、TAE、PAE、EAE、PAD、AEV、AEE或ADV。在一些实施方案中,[N2]包含YPAE(SEQ ID NO:21)、YQAE(SEQ ID NO:25)、YTAE(SEQ ID NO:24)、TEAE(SEQ ID NO:587)、YPAD(SEQ ID NO:588)、QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)、EAEV(SEQ ID NO:590)或PADV(SEQ ID NO:19)。在一些实施方案中,[N2]为或包含YPAEV(SEQ ID NO:1)、YQAEV(SEQ ID NO:6)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、TEAEV(SEQ ID NO:12)或YPADV(SEQ ID NO:13)。在一些实施方案中,[N2]为YPAEV(SEQ ID NO:1)。在一些实施方案中,[N2]-[N3]包含AEVVQK(SEQ ID NO:36)、AEEVQK(SEQ ID NO:39)、AEVVQN(SEQ ID NO:591)或ADVVQK(SEQ ID NO:593)。在一些实施方案中,[N2]-[N3]包含PAEVVQN(SEQ ID NO:594)、QAEVVQK(SEQ ID NO:52)、TAEVVQK(SEQ ID NO:49)、PAEVVQK(SEQ ID NO:20)、PAEEVQK(SEQ ID NO:51)、EAEVVQK(SEQ ID NO:595)或PADVVQK(SEQ ID NO:596)。在一些实施方案中,[N2]-[N3]为或包含YPAEVVQK(SEQ ID NO:943)、YQAEVVQK(SEQ ID NO:951)、YTAEVVQK(SEQ ID NO:948)、YPAEEVQK(SEQ ID NO:950)、YPAEVVQN(SEQ ID NO:964)、TEAEVVQK(SEQ ID NO:965)或YPADVVQK(SEQ ID NO:966);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N2]-[N3]为YPAEVVQK(SEQ ID NO:943)。

[1139] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的肽进一步包含[N1],其包含位置X_D、X_E和X_F。在一些实施方案中,[N1]的位置X_D为Q或S。在一些实施方案中,[N1]的位置X_E为S、L、A或T。在一些实施方案中,[N1]的位置X_F为S、Y或T。在一些实施方案中,[N1]包含QS、SL、SA、QT、LS、LY、AT、TS或SS。在一些实施方案中,[N1]为或包含QSS、SLS、SLY、SAT或QTS。在一些实施方案中,[N1]为QSS。在一些实施方案中,[N1]-[N2]包含SSYPA(SEQ ID NO:63)、LSYQA(SEQ ID NO:597)、LSYTA(SEQ ID NO:598)、LYYPA(SEQ ID NO:600)、ATYPA(SEQ ID NO:601)、LSYPA(SEQ ID NO:603)或TSTEA(SEQ ID NO:605)。在一些实施方案中,[N1]-[N2]包含SSYPAE(SEQ ID NO:79)、LSYQAE(SEQ ID NO:607)、LSYTAE(SEQ ID NO:610)、LYYPAE(SEQ ID NO:611)、ATYPAE(SEQ ID NO:613)、LSYPAE(SEQ ID NO:616)、TSTEAE(SEQ ID NO:619)或LSYPAD(SEQ ID NO:621)。在一些实施方案中,[N1]-[N2]为或包含QSSYPAEV(SEQ ID NO:96)、SLSYQAEV(SEQ ID NO:622)、SLSYTAEV(SEQ ID NO:623)、SLYYPAEV(SEQ ID NO:624)、SATYPAEV(SEQ ID NO:625)、SLSYPAEV(SEQ ID NO:629)、SLSYPAEE(SEQ ID NO:632)、QTSTEAEV(SEQ ID NO:633)或SLSYPADV(SEQ ID NO:634);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]为QSSYPAEV(SEQ ID

NO:96)。在一些实施方案中,[N1]-[N2]-[N3]为或包含QSSYPAEVVQK(SEQ ID NO:150)、SLSYQAEVVQK(SEQ ID NO:635)、SLSYTAEVVQK(SEQ ID NO:637)、SLYYPAEVVQK(SEQ ID NO:639)、SATYPAEVVQK(SEQ ID NO:641)、SLSYPAEVVQK(SEQ ID NO:642)、SLSYPAEEVQK(SEQ ID NO:643)、SLSYPAEVVQN(SEQ ID NO:644)、QTSTEAEVVQK(SEQ ID NO:645)或SLSYPADVVQK(SEQ ID NO:646);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]-[N3]为QSSYPAEVVQK(SEQ ID NO:150)。

[1140] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的肽进一步包含[N0],其中[N0]包含位置X_A、X_B和X_C。在一些实施方案中,[N0]的位置X_A为T。在一些实施方案中,[N0]的位置X_B为N。在一些实施方案中,[N0]的位置X_C为N、T、S或K。在一些实施方案中,[N0]包含TN、NS、NT、NN或NK。在一些实施方案中,[N0]为或包含TNS、TNT、TNN或TNK。在一些实施方案中,[N0]为TNN。在一些实施方案中,[N0]-[N1]为或包含TNNQSS(SEQ ID NO:183)、TNSSLS(SEQ ID NO:647)、TNSSLY(SEQ ID NO:648)、TNTSAT(SEQ ID NO:649)、TNNQTS(SEQ ID NO:650)或TNKSAT(SEQ ID NO:651);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[N0]-[N1]-[N2]-[N3]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNSSLSYQAEVVQK(SEQ ID NO:652)、TNSSLSYTAEVVQK(SEQ ID NO:654)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNTSATYPAEVVQK(SEQ ID NO:656)、TNSSLSYPAEVVQK(SEQ ID NO:657)、TNSSLSYPAEEVQK(SEQ ID NO:658)、TNSSLSYPAEVVQN(SEQ ID NO:660)、TNNQTSTEAEVVQK(SEQ ID NO:662)、TNKSATYPAEVVQK(SEQ ID NO:663)或TNSSLSYPADVVQK(SEQ ID NO:665);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]-[N2]-[N3]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1141] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的肽进一步包含[N4],其包含位置X_G和X_H。在一些实施方案中,[N4]的位置X_G为T、P或N。在一些实施方案中,[N4]的位置X_H为A或D。在一些实施方案中,[N4]为或包含TA、TD、PA或NA。在一些实施方案中,[N4]为TA。在一些实施方案中,[N3]-[N4]为或包含VQKTA(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKKTA(SEQ ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SEQ ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571);包含任何上述氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四

个不同氨基酸的氨基酸序列。在一些实施方案中,[N3]-[N4]为VQKTA(SEQ ID NO:564)。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNSSLSYQAEV VQKTA(SEQ ID NO:2064)、TNSSLSYTAEVVQKTA(SEQ ID NO:2065)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNTSAT YPAEVVQKTA(SEQ ID NO:2067)、TNSSLSYPAEVVQKTA(SE Q ID NO:2068)、TNSSLSYPAEEVQKTA(SEQ ID NO:2069)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQNT A(SEQ ID NO:2071)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)、TNNQTSTEAEVVQKTA(SEQ ID NO:2074)、TNKSATYPAEVVQKTA(SEQ ID NO:2075)或TNSSLSYPADVVQKTA(SEQ ID NO:2076);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1142] 在一些实施方案中,[N1]紧接在[N0]之后存在。在一些实施方案中,[N2]紧接在[N1]之后存在。在一些实施方案中,[N3]紧接在[N2]之后存在。在一些实施方案中,[N4]紧接在[N3]之后存在。在一些实施方案中,所述肽自N端至C端包含[N2]-[N3]。在一些实施方案中,所述肽自N端至C端包含[N1]-[N2]-[N3]。在一些实施方案中,所述肽自N端至C端包含[N1]-[N2]-[N3]-[N4]。在一些实施方案中,所述肽自N端至C端包含[N0]-[N1]-[N2]-[N3]。在一些实施方案中,所述肽自N端至C端包含[N0]-[N1]-[N2]-[N3]-[N4]。

[1143] 在一些实施方案中,本文所述的肽包含具有式[B]-[C]的氨基酸序列,其中[B]包含位置X1、X2和X3,且[C]包含氨基酸序列YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,[B]X1的位置X1为Q、T、S、A、I、L或H。在一些实施方案中,[B]X1的位置X1为Q、T、S、A或H。在一些实施方案中,[B]的位置X2为S、G或A。在一些实施方案中,[B]的位置X2为S或G。在一些实施方案中,[B]的位置X3为S、K、L、R或A。在一些实施方案中,[B]的位置X3为S、K、L或R。在一些实施方案中,[B]包含位置X1处的Q。在一些实施方案中,[B]包含位置X2处的S。在一些实施方案中,[B]包含位置X3处的S。在一些实施方案中,[B]包含QS、TS、SS、AG、IG、QA、AS、LG、HS、SK、SL、SR、GA或GS。在一些实施方案中,[B]为或包含QSS、TSL、SSS、QSR、QSK、AGA、IGS、QAS、ASS、LGS或HSS。在一些实施方案中,[B]的氨基酸序列为QSS。在一些实施方案中,[B]-[C]包含SSYPAEVVQK(SEQ ID NO:572)、SKYPAEVVQK(SEQ ID NO:573)、SLYPAEVVQK(SEQ ID NO:574)、SRYPAEVVQK(SEQ ID NO:575)、GAYPAEVVQK(SEQ ID NO:576)、GSYPAEVVQK(SEQ ID NO:580)或ASYPAEVVQK(SEQ ID NO:582)。在一些实施方案中,[B]-[C]为或包含QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)或HSSYPAEVVQK(SEQ ID NO:162);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一

个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[B]-[C]为QSSYPAEVVQK(SEQ ID NO:150)。

[1144] 在一些实施方案中,包含式[B]-[C]的肽进一步包含[A],其包含位置X_A、X_B和X_C。在一些实施方案中,[A]的位置X_A为T、I或N。在一些实施方案中,[A]的位置X_B为N。在一些实施方案中,[A]的位置X_C为N、T、S或K。在一些实施方案中,[A]包含TN、IN、NN、NT、NS或NK。在一些实施方案中,[A]为或包含TNN、TNT、INN、NNN、TNS或TNK。在一些实施方案中,[A]为TNN。在一些实施方案中,[A]-[B]为或包含TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNHSS(SEQ ID NO:194)、INNQSS(SEQ ID NO:196)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)或TNKQAS(SEQ ID NO:201);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]的氨基酸序列为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[A]-[B]-[C]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVVQK(SEQ ID NO:510)、TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQK(SEQ ID NO:517)、TNTASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNHSSYPAEVVQK(SEQ ID NO:525)、INNQSSYPAEVVQK(SEQ ID NO:543)、NNNQSRYPAEVVQK(SEQ ID NO:552)、TNSTSLYPAEVVQK(SEQ ID NO:556)或TNKQASYPAEVVQK(SEQ ID NO:563);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[A]-[B]-[C]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1145] 在一些实施方案中,包含式[B]-[C]的肽进一步包含[D],其中[D]包含位置X4和X5。在一些实施方案中,[D]的位置X4为T或N。在一些实施方案中,[D]的位置X5为A。在一些实施方案中,[D]为或包含TA或PA。在一些实施方案中,[D]的氨基酸序列为TA。在一些实施方案中,[C]-[D]为或包含YPAEVVQKTA(SEQ ID NO:584)或YPAEVVQKPA(SEQ ID NO:586);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[C]-[D]的氨基酸序列为YPAEVVQKTA(SEQ ID NO:584)。在一些实施方案中,[A]-[B]-[C]-[D]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEV VQKTA(SEQ ID NO:1539)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)、TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNI

GSYPAEVVQKTA(SEQ ID NO:1112)、TNNQASYPAEVVQKTA(SEQ ID NO:1194)、TNTASSYPAEVVQKTA(SEQ ID NO:1575)、TNNLGSYPAEVVQKTA(SEQ ID NO:1027)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNNSSSYPAEVVQKPA(SEQ ID NO:1142)、INNQSSYPAEVVQKTA(SEQ ID NO:1024)、NNNQS RYPAEVVQKTA(SEQ ID NO:1601)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)或TNKQASYPAEVVQKTA(SEQ ID NO:1587);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[C]-[D]为YPAE VVQKTA(SEQ ID NO:584)。在一些实施方案中,[A]-[B]-[C]-[D]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1146] 在一些实施方案中,本文所述的肽包含具有式[B]-[C]的氨基酸序列,其中[B]包含位置X1、X2和X3,且[C]包含氨基酸序列YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,[B]的位置X1为Q或S。在一些实施方案中,[B]的位置X2为S、L或A。在一些实施方案中,[B]的位置X3为S、Y或T。在一些实施方案中,[B]包含位置X1处的Q。在一些实施方案中,[B]包含位置X2处的S。在一些实施方案中,[B]包含位置X3处的S。在一些实施方案中,[B]包含QS、SL、SA、LY、AT、LS或SS。在一些实施方案中,[B]为或包含QSS、SLY、SAT或SLS。在一些实施方案中,[B]为QSS。在一些实施方案中,[B]-[C]包含SSYPAEVVQK(SEQ ID NO:572)、LYYPAEVVQK(SEQ ID NO:702)、ATYPAEVVQK(SEQ ID NO:718)或LSYPAEVVQK(SEQ ID NO:703)。在一些实施方案中,[B]-[C]为或包含QSSYPAEVVQK(SEQ ID NO:150)、SLYYPAEVVQK(SEQ ID NO:639)、SATYPAEVVQK(SEQ ID NO:641)或SLSYPAEVVQK(SEQ ID NO:642);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[B]-[C]为QSSYPAEVVQK(SEQ ID NO:150)。

[1147] 在一些实施方案中,包含式[B]-[C]的肽进一步包含[A],其包含位置X_A、X_B和X_C。在一些实施方案中,[A]的位置X_A为T。在一些实施方案中,[A]的位置X_B为N。在一些实施方案中,[A]的位置X_C为N、T、S或K。在一些实施方案中,[A]包含TN、NS、NT、NK或NN。在一些实施方案中,[A]为或包含TNN、TNS、TNT或TNK。在一些实施方案中,[A]的氨基酸序列为TNN。在一些实施方案中,[A]-[B]为或包含TNNQSS(SEQ ID NO:183)、TNSSLY(SEQ ID NO:648)、TNTSAT(SEQ ID NO:649)、TNSSLS(SEQ ID NO:647)或TNKSAT(SEQ ID NO:651);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[A]-[B]-[C]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNTSATYPAEVVQK(SEQ ID NO:656)、TNSSLSYPAEVVQK(SEQ ID NO:657)或TNKSATYPAEVVQK(SEQ ID NO:663);包含任何上述氨基酸序列的任何部

分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[A]-[B]-[C]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1148] 在一些实施方案中,包含式[B]-[C]的肽进一步包含[D],其中[D]包含位置X4和X5。在一些实施方案中,[D]的位置X4为T、N或P。在一些实施方案中,[D]的位置X5为A或D。在一些实施方案中,[D]为或包含TA、TD、NA或PA。在一些实施方案中,[D]的氨基酸序列为TA。在一些实施方案中,[C]-[D]为或包含YPAEVVQKTA(SE Q ID NO:584)、YPAEVVQKTD(SEQ ID NO:719)、YPAEVVQK NA(SEQ ID NO:724)或YPAEVVQKPA(SEQ ID NO:586);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[C]-[D]为YPAEVVQKTA(SEQ ID NO:584)。在一些实施方案中,[A]-[B]-[C]-[D]为或包含TNNQSSYPAEVV QKTA(SEQ ID NO:1533)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNTSATYPAEVVQKTA(SEQ ID NO:2067)、TNSSLSYP AEVVQKTA(SEQ ID NO:2068)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)或TNKSATYPAEVVQKT A(SEQ ID NO:2075);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]-[C]-[D]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1149] 在一些实施方案中,[B]紧接在[A]之后存在。在一些实施方案中,[C]紧接在[B]之后存在。在一些实施方案中,[D]紧接在[C]之后存在。在一些实施方案中,所述肽自N端至C端包含[B]-[C]。在一些实施方案中,所述肽自N端至C端包含[B]-[C]。在一些实施方案中,所述肽自N端至C端包含[A]-[B]-[C]。在一些实施方案中,所述肽自N端至C端包含[B]-[C]-[D]。在一些实施方案中,所述肽自N端至C端包含[A]-[B]-[C]-[D]。

[1150] 在一些实施方案中,本文所述的肽包含式[K1]-[K2],其中[K1]包含LSY或LYY,且[K2]包含位置X1、X2、X3和X4。在一些实施方案中,[K1]包含LSY。在一些实施方案中,[K2]的位置X1为Q、T或P。在一些实施方案中,[K2]的位置X2为A,在一些实施方案中,[K2]的位置X3为E或D。在一些实施方案中,[K2]的位置X4为V或E。在一些实施方案中,[K2]包含QA、TA、PA、EV、EE、DV、AE或AD。在一些实施方案中,[K2]包含QAE、TAE、PAE、PAD、AEV、AEE或ADV。在一些实施方案中,[K2]为或包含QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)或PADV(SEQ ID NO:19)。在一些实施方案中,[K1]-[K2]包含LSYQA(SEQ ID NO:597)、LSYTA(SEQ ID NO:598)、LYYPA(SEQ ID NO:600)或LSYPA(SEQ ID NO:603)。在一些实施方案中,[K1]-[K2]包含LSYQAE(SEQ ID NO:607)、LSYTAE(SEQ ID NO:610)、LYYPAE(SEQ ID NO:611)、LSYPAE(SEQ ID NO:616)或LSYPAD(SEQ ID NO:621)。在一

些实施方案中,[K1]-[K2]为或包含LSYQAEV(SEQ ID NO:667)、LSYTAEV(SEQ ID NO:668)、LYYPAEV(SEQ ID NO:669)、LSYPAEV(SEQ ID NO:671)、LSYPAEE(SEQ ID NO:673)或LSYPADV(SEQ ID NO:674);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1151] 在一些实施方案中,包括包含式[K1]-[K2]的氨基酸序列的肽进一步包含[K0],其包含TNNS(SEQ ID NO:14)。在一些实施方案中,[K0]-[K1]包含TNSSLS(SEQ ID NO:647)或TNSSLY(SEQ ID NO:648)。在一些实施方案中,[K0]-[K1]为或包含TNSSLSY(SEQ ID NO:676)或TNSSLYY(SEQ ID NO:678);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K0]-[K1]-[K2]包含TNSSLSYQA(SEQ ID NO:679)、TNSSLSYTA(SEQ ID NO:681)、TNSSLYYPA(SEQ ID NO:682)或TNSSLSYPA(SEQ ID NO:683)。在一些实施方案中,[K0]-[K1]-[K2]包含TNSSLSYQAE(SEQ ID NO:684)、TNSSLSYTAE(SEQ ID NO:685)、TNSSLYYPAE(SEQ ID NO:686)、TNSSLSYPAE(SEQ ID NO:687)或TNSSLSYPAD(SEQ ID NO:689)。在一些实施方案中,[K0]-[K1]-[K2]为或包含TNSSLSYQAEV(SEQ ID NO:692)、TNSSLSYTAEV(SEQ ID NO:693)、TNSSLYYPAEV(SEQ ID NO:696)、TNSSLSYPAEV(SEQ ID NO:697)、TNSSLSYPAEE(SEQ ID NO:698)或TNSSLSYPADV(SEQ ID NO:699);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1152] 在一些实施方案中,包括包含式[K1]-[K2]的氨基酸序列的肽进一步包含[K3],其中[K3]包含位置X_A、X_B和X_C。在一些实施方案中,[K3]的位置X_A为V。在一些实施方案中,[K3]的位置X_B为Q。在一些实施方案中,[K3]的位置X_C为K或N。在一些实施方案中,[K3]包含VQ、QK或QN。在一些实施方案中,[K3]为或包含VQK或VQN。在一些实施方案中,K2]-[K3]为或包含QAEVVQK(SEQ ID NO:52)、TAEVVQK(SEQ ID NO:49)、PAEVVQK(SEQ ID NO:20)、PAEEVQK(SEQ ID NO:51)、PAEVVQN(SEQ ID NO:594)或PADVVQK(SEQ ID NO:596);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K1]-[K2]-[K3]为或包含LSYQAEVVQK(SEQ ID NO:700)、LSYTAEVVQK(SEQ ID NO:701)、LYYPAEVVQK(SEQ ID NO:702)、LSYPAEVVQK(SEQ ID NO:703)、LSYPAEEVQK(SEQ ID NO:704)、LSYPAEVVQN(SEQ ID NO:706)或LSYPADVVQK(SEQ ID NO:708);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的

氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K0]-[K1]-[K2]-[K3]为或包含TNSSLSYQAEVVQK(SEQ ID NO:652)、TNSSLSYTAEVVQK(SEQ ID NO:654)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNSSLSYPAEVVQK(SEQ ID NO:657)、TNSSLSYPAEEVQK(SEQ ID NO:658)、TNSSLSYPAEVVQN(SEQ ID NO:660)或TNSSLSYPADVVQK(SEQ ID NO:665);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1153] 在一些实施方案中,包括包含式[K1]-[K2]的氨基酸序列的肽进一步包含[K4],其中[K4]包含位置X_D和X_E。在一些实施方案中,[K4]的位置X_D为T、P或N。在一些实施方案中,[K4]的位置X_E为A或D。在一些实施方案中,[K4]为或包含TA、TD、PA或NA。在一些实施方案中,[K3]-[K4]为或包含VQKTA(SEQ ID NO:564)、VQK TD(SEQ ID NO:714)、VQNTA(SEQ ID NO:715)、VQKNA(SEQ ID NO:570)或VQKPA(SEQ ID NO:567);包含任何上述氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为或包含TNSSLSYQAEVVQKTA(SEQ ID NO:2064)、TNSSLSYTAEVVQKTA(SEQ ID NO:2065)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNSSLSYPAEVVQKTA(SEQ ID NO:2068)、TNSSLSYPAEEVQKTA(SEQ ID NO:2069)、TNSSLS YPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQNTA(SE Q ID NO:2071)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)或TNSSLSYPADVVQK TA(SEQ ID NO:2076);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1154] 在一些实施方案中,[K2]紧接在[K1]之后存在。在一些实施方案中,[K1]紧接在[K0]之后存在。在一些实施方案中,[K3]紧接在[K2]之后存在。在一些实施方案中,[K4]紧接在[K3]之后存在。在一些实施方案中,所述肽自N端至C端包含[K1]-[K2]。在一些实施方案中,所述肽自N端至C端包含[K0]-[K1]-[K2]。在一些实施方案中,所述肽自N端至C端包含[K1]-[K2]-[K3]。在一些实施方案中,所述肽自N端至C端包含[K0]-[K1]-[K2]-[K3]。在一些实施方案中,所述肽自N端至C端包含[K1]-[K2]-[K3]-[K4]。在一些实施方案中,所述肽自N端至C端包含[K0]-[K1]-[K2]-[K3]-[K4]。

[1155] 在一些实施方案中,本文所述的肽包含氨基酸序列,所述氨基酸序列包含来自表1、2A、2B、9和15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列包含来自SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中

任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列包含来自SEQ ID NO:943或2064-2080中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸。在一些实施方案中,所述肽包含来自SEQ ID NO:943或946-966中任一个的至少3、4、5、6或7个连续氨基酸。

[1156] 在一些实施方案中,3个连续氨基酸包含YPA。在一些实施方案中,4个连续氨基酸包含YPAE(SEQ ID NO:21)。在一些实施方案中,5个连续氨基酸包含YPAEV(SEQ ID NO:1)。在一些实施方案中,6个连续氨基酸包含YPAEVV(SEQ ID NO:725)。在一些实施方案中,7个连续氨基酸包含YPAEVVQ(SEQ ID NO:726)。在一些实施方案中,所述氨基酸序列包含YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,所述氨基酸序列由YPAEVVQK(SEQ ID NO:943)组成。

[1157] 在一些实施方案中,本文所述的肽包含氨基酸序列,所述氨基酸序列相对于表1、2A、2B、9和15-20中所提供的任一序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于表1、2A、2B、9和15-20中所提供的任一序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸。

[1158] 在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸。

[1159] 在一些实施方案中,所述肽包含胺基序列,所述胺基序列相对于SEQ ID NO:2024-2063中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于SEQ ID NO:2024-2063中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸。在一些实施方案中,相对于SEQ ID NO:2024-2063中任一个的氨基酸序列包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列的不同氨基酸存在于以下位置中的一处或多处:(i)位置1,其中不同氨基酸为T或L;(ii)位置2,其中不同氨基酸为N、L、K、A、T

或P;(iii)位置3,其中不同氨基酸为N、K、L、A、Y或S;(iv)位置4,其中不同氨基酸为Q、L、T、S、F、Y、K或A;(v)位置5,其中不同氨基酸为S、H、A、M、Q、T、V或F;(vi)位置6,其中不同氨基酸为S、P、V、A、Q、L、T、N或M;(vii)位置7,其中不同氨基酸为Y、H、S、V、A、L或T;(viii)位置8,其中不同氨基酸为D、P、A、Q、F、L、S、H或M;(ix)位置9,其中不同氨基酸为F、A、L、D或Q;(x)位置10,其中不同氨基酸为T、E、I或S;(xi)位置11,其中不同氨基酸为V、A、N或S;(xii)位置12,其中不同氨基酸为V、L或P;(xiii)位置13,其中不同氨基酸为Q、E或P;(xiv)位置14,其中不同氨基酸为K、N、S或L;(xv)位置15,其中不同氨基酸为T、V、M或L;和/或(xvi)位置16,其中不同氨基酸为A、G或R。

[1160] 在一些实施方案中,所述肽包含胺基序列,所述胺基序列相对于SEQ ID NO:1632-2023中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,所述肽包含氨基酸序列,所述氨基酸序列相对于SEQ ID NO:1632-2023中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸。在一些实施方案中,相对于SEQ ID NO:1623-2023中任一个的氨基酸序列包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列的不同氨基酸存在于以下位置中的一处或多处:(i)位置1,其中不同氨基酸为T、G、N、S、E、L、Y、V或I;(ii)位置2,其中不同氨基酸为D、N、K、E、V、G、R、L、H、F、P、T、A、S、I或Y;(iii)位置3,其中不同氨基酸为Y、N、K、T、W、Q、M、V、C、A、L、F、H、G、R、S或P;(iv)位置4,其中不同氨基酸为H、Q、P、E、R、K、A、S、V、L、T、D、I、G、M或N;(v)位置5,其中不同氨基酸为R、S、K、N、H、G、W、A、P、V、Q、Y、L或F;(vi)位置6,其中不同氨基酸为G、S、F、R、W、H、I、C、M、A、Y、K、N、Q、V、P、E、D、T或L;(vii)位置7,其中不同氨基酸为D、Y、S、I、H、F、P、K、R、G、L、Q、A、M、T、N、V、W、C或E;(viii)位置8,其中不同氨基酸为P、L、Q、T、W、V、G、K、I、Y、N、H、R、D、S、M、A、F或E;(ix)位置9,其中不同氨基酸为A、R、T、Q、S、M、L、E、K、V、G、D、N、H、F、P或I;(x)位置10,其中不同氨基酸为K、E、Q、H、V、G、R、S、P、I、N、M、A、L、D或T;(xi)位置11,其中不同氨基酸为V、A、E、N、R、L、M、T、Q、S、K、C、G、D、Y、P、H、F或I;(xii)位置12,其中不同氨基酸为V、P、L、S、T、N、A、G、K、R、I、H、E、Q或M;(xiii)位置13,其中不同氨基酸为Q、K、N、A、H、R、T、V、E、I、P、G、S或L;(xiv)位置14,其中不同氨基酸为K、E、I、Y、Q、R、G、D、L、N或S;(xv)位置15,其中不同氨基酸为S、T、N、Q、I、P、E、G、K、M或H;和/或(xvi)位置16,其中不同氨基酸为A、D、L、Y、Q或T。

[1161] 在一些实施方案中,所述肽包含表1、2A、2B、9和15-20中所提供的任何序列的氨基酸序列。在一些实施方案中,所述肽包含SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列。在一些实施方案中,所述肽包含SEQ ID NO:943或2064-2080中任一个的氨基酸序列。在一些实施方案中,所述肽包含SEQ ID NO:2024-2063中任一个的氨基酸序列。在一些实施方案中,所述肽包含SEQ ID NO:1632-2023中任一个的氨基酸序列。在一些实施方案中,所述肽包含SEQ ID NO:943的氨基酸序列。

[1162] 在一些实施方案中,所述肽包含由本文所述的核苷酸序列(例如,表2A的核苷酸序列)编码的氨基酸序列。在一些实施方案中,所述肽包含由以下编码的氨基酸序列:相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些

实施方案中,所述肽包含由以下编码的氨基酸序列:相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。在一些实施方案中,所述肽包含由以下编码的氨基酸序列:SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。

[1163] 在一些实施方案中,编码本文所述的肽的核苷酸序列包含本文所述的核苷酸序列,例如,如表2A中所述。在一些实施方案中,编码本文所述的肽的核苷酸序列为密码子优化的。在一些实施方案中,编码本文所述的肽的核苷酸序列为分离的,例如重组的。

[1164] 在一些实施方案中,编码本文所述的肽的核苷酸序列包含SEQ ID NO:944的核苷酸序列,或相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,编码本文所述的肽的核苷酸序列包含相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。在一些实施方案中,编码本文所述的肽的核酸包含核苷酸序列,所述核苷酸序列包含SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。

[1165] 在一些实施方案中,本文所述的肽与活性剂融合或偶合,例如结合。在一些实施方案中,活性剂为治疗剂。在一些实施方案中,活性剂包含治疗性蛋白质、抗体分子、酶、基因组编辑系统的一种或多种组分、与治疗剂融合或偶合(例如,共价或非共价)的Fc多肽和/或RNAi剂(例如,dsRNA、反义寡核苷酸(ASO)、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA)。在一些实施方案中,治疗剂为抗体。在一些实施方案中,本文所述的肽与抗体的Fc区,例如在Fc区的C端或Fc区的N端融合或偶合,例如结合(例如,直接地或间接地)。在一些实施方案中,治疗剂为RNAi剂。在一些实施方案中,RNAi剂为siRNA或ASO。在一些实施方案中,ASO或siRNA包含至少一个(例如,一个或多个或所有)经修饰的核苷酸。在一些实施方案中,本文所述的肽与RNAi剂的至少一条链融合或偶合,例如结合(例如,直接地或经由接头间接地)。在一些实施方案中,本文所述的肽与RNAi剂的至少一条链的C端结合(例如,直接地或经由接头间接地)。在一些实施方案中,本文所述的肽与RNAi剂的至少一条链的内部核苷酸结合(例如,直接地或经由接头间接地)。在一些实施方案中,至少一条链为有义链。在一些实施方案中,治疗剂调节(例如,抑制、减少或增加)CNS相关基因、mRNA和/或蛋白质的表达。

[1166] 在一些实施方案中,活性剂为诊断剂。在一些实施方案中,诊断剂为或包含成像剂(例如,与可检测部分偶合的蛋白质或小分子化合物)。在一些实施方案中,成像剂包含PET或MRI配位体,或与可检测部分偶合的抗体分子。在一些实施方案中,可检测部分为或包含放射性标记、荧光团、发色团或亲和标签。在一些实施方案中,放射性标记为或包含tc99m、碘-123、自旋标记、碘-131、铟-111、氟-19、碳-13、氮-15、氧-17、钆、锰或铁。在一些实施方案中,活性剂为小分子。在一些实施方案中,活性剂为核糖核酸复合物(例如,Cas9/gRNA复合物)、质体、闭端DNA、circ-RNA或mRNA。

[1167] 在一些实施方案中,至少1-5个(例如,至少1、2、3、4或5个)肽与活性剂(例如,治疗剂或诊断剂)融合或偶合,例如结合。在一些实施方案中,至少1-5个(例如,至少1、2、3、4或5

个)肽包含相同氨基酸序列。在一些实施方案中,至少1-5个(例如,至少1、2、3、4或5个)肽包含不同氨基酸序列。在一些实施方案中,至少1-5个(例如,至少1、2、3、4或5个)肽以串联形式(例如,直接地或经由接头间接地连接)或以多聚体组态存在。在一些实施方案中,所述肽包含至少3、4、5、6、7、8、9、10、11、12、15、20、25、30或35个氨基酸的长度的氨基酸序列。

[1168] 在一些实施方案中,所述肽与活性剂共价连接(例如,直接地或经由接头间接地)。在一些实施方案中,所述肽经由接头与活性剂结合。在一些实施方案中,接头为可裂解接头或不可裂解接头。在一些实施方案中,可裂解接头为pH敏感性接头或酶敏感性接头。在一些实施方案中,pH敏感性接头包含肼/腙接头或二硫化物接头。在一些实施方案中,酶敏感性接头包含基于肽的接头,例如对蛋白酶(例如,溶酶体蛋白酶)敏感的肽接头;或β-葡萄糖醛酸苷接头。在一些实施方案中,不可裂解接头为包含硫醚基或马来酰亚胺己酰基的接头。在一些实施方案中,肽和活性剂在转译后融合或偶合,例如使用点击化学。在一些实施方案中,肽和活性剂经由化学诱导的二聚作用融合或偶合。在一些实施方案中,肽相对于活性剂存在于N端。在一些实施方案中,肽相对于活性剂存在于C端。

[1169] 在一些实施方案中,肽存在于载剂上或与载剂偶合。在一些实施方案中,载剂包含外来体、微囊泡或脂质纳米颗粒(LNP)。在一些实施方案中,载剂包含治疗剂(例如,RNAi剂(例如,dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA、反义寡核苷酸剂(ASO)或snoRNA)、mRNA、核糖核蛋白复合物(例如,Cas9/gRNA复合物)或circRNA)。在一些实施方案中,所述肽存在于载剂之表面上。在一些实施方案中,载剂之表面的至少10%、20%、30%、40%、50%、60%、70%或80%包含至少1-5个(例如,至少1、2、3、4或5个)本文所述的肽。

[1170] 本公开还提供编码本文所述的任何肽的核酸或多核苷酸,以及包含其的AAV衣壳变体、AAV颗粒、载体和细胞。

[1171] AAV衣壳变体

[1172] 在一些实施方案中,本文所述的AAV衣壳变体(例如,AAV5衣壳变体)包含相对于SEQ ID NO:138编号的位置577处非T的氨基酸(例如,Y、N或C)。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:138编号的位置577处的Y。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:138编号的位置577处的N。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:138编号的位置577处的C。

[1173] 在一些实施方案中,本文所述的AAV衣壳变体(例如,AAV5衣壳变体)包含超过一个置换相对于SEQ ID NO:138编号的位置577处的苏氨酸(T)的氨基酸。在一些实施方案中,两个、三个、四个、五个、六个、七个、八个、九个或十个氨基酸的插入物置换相对于SEQ ID NO:138编号的位置577处的T。在一些实施方案中,八个氨基酸的插入物置换相对于SEQ ID NO:138编号的位置577处的T。

[1174] 在一些实施方案中,本文所述的AAV颗粒包含AAV衣壳变体,例如本文所述的AAV衣壳变体(例如,包含本文所述的肽或氨基酸序列的AAV衣壳变体)。在一些实施方案中,AAV衣壳变体包含如表1、2A、2B、9或15-20中任一个所陈述的肽。

[1175] 在一些实施方案中,本文所述的AAV衣壳变体包含具有式[N2]-[N3]的氨基酸序列,其中[N2]包含位置X1、X2、X3、X4和X5且[N3]包含VQK、VQN、EQK、VKK、VHK、VQQ或LQK的氨基酸序列。在一些实施方案中,[N2]的位置X1为Y、N、C或T。在一些实施方案中,[N2]的位置

X2为P、E、K、T或Q。在一些实施方案中,[N2]的位置X3为A或P。在一些实施方案中,[N2]的位置X4为E、S、D或A。在一些实施方案中,[N2]的位置X5为V、L或E。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N3]的氨基酸序列包含VQK。在一些实施方案中,[N3]的氨基酸序列为VQK。

[1176] 在一些实施方案中,本文所述的AAV衣壳变体包含具有式[N2]-[N3]的氨基酸序列,其中[N2]包含位置X1、X2、X3、X4和X5且[N3]包含VQK、EQK、VKK、VHK、VQQ或LQK的氨基酸序列。在一些实施方案中,[N3]包含VQK、EQK或VKK的氨基酸序列。在一些实施方案中,[N3]包含氨基酸序列VQK。在一些实施方案中,[N2]的位置X1为Y、N或C。在一些实施方案中,[N2]的位置X1为Y或N。在一些实施方案中,[N2]的位置X2为P、K、T或Q。在一些实施方案中,[N2]的位置X2为P、T或Q。在一些实施方案中,[N2]的位置X3为A或P。在一些实施方案中,[N2]的位置X3为A。在一些实施方案中,[N2]的位置X4为E、S或A。在一些实施方案中,[N2]的位置X5为V、L或E。在一些实施方案中,[N2]的位置X5为V或L。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N2]包含YPA、YPP、NKA、YTA、YQA、YTP、NPA、CPA、THA、PAE、PPS、KAE、TAE、QAE、TPS、PAA、HAS、AEV、PSL、AEE或AAV。在一些实施方案中,[N2]包含YPAE(SEQ ID NO:21)、YPPS(SEQ ID NO:22)、NKAE(SEQ ID NO:23)、YTAE(SEQ ID NO:24)、YQAE(SEQ ID NO:25)、YTPS(SEQ ID NO:26)、YPAA(SEQ ID NO:27)、NPAE(SEQ ID NO:28)、CPAE(SEQ ID NO:29)、THAS(SEQ ID NO:30)、PAEV(SEQ ID NO:17)、PPSL(SEQ ID NO:31)、KAEV(SEQ ID NO:32)、TAEV(SEQ ID NO:16)、PAEE(SEQ ID NO:18)、QAEV(SEQ ID NO:15)、TPSL(SEQ ID NO:33)、PAAV(SEQ ID NO:34)或QAEE(SEQ ID NO:35)。在一些实施方案中,[N2]为或包含YPAEV(SEQ ID NO:1)、YPPSL(SEQ ID NO:2)、NKAEV(SEQ ID NO:3)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、YQAEV(SEQ ID NO:6)、YTPSL(SEQ ID NO:7)、YPAAV(SEQ ID NO:8)、NPAEV(SEQ ID NO:9)、CPAEV(SEQ ID NO:10)或YQAEE(SEQ ID NO:11)。在一些实施方案中,[N2]为YPAEV(SEQ ID NO:1)。在一些实施方案中,[N2]-[N3]包含AEVVQK(SEQ ID NO:36)、PSLVQK(SEQ ID NO:37)、AEVEQK(SEQ ID NO:38)、AEEVQK(SEQ ID NO:39)、PSLEQK(SEQ ID NO:40)、PSLVKK(SEQ ID NO:41)、AEVVKK(SEQ ID NO:42)、AEVVHK(SEQ ID NO:43)、AAVVQK(SEQ ID NO:44)、AEVVQQ(SEQ ID NO:45)或AEVLQK(SEQ ID NO:46)的氨基酸序列。在一些实施方案中,[N2]-[N3]包含氨基酸序列PAEVVQK(SEQ ID NO:20)、PPSLVQK(SEQ ID NO:47)、KAEVVQK(SEQ ID NO:48)、TAEVVQK(SEQ ID NO:49)、PAEVEQK(SEQ ID NO:50)、PAEEVQK(SEQ ID NO:51)、QAEVVQK(SEQ ID NO:52)、TPSLVQK(SEQ ID NO:53)、PPSLEQK(SEQ ID NO:54)、PPSLVKK(SEQ ID NO:55)、PAEVVKK(SEQ ID NO:56)、PAEVVHK(SEQ ID NO:57)、PAAVVQK(SEQ ID NO:58)、PAEVVQQ(SEQ ID NO:59)、TAEVVKK(SEQ ID NO:60)、PAEVLQK(SEQ ID NO:61)或QAEEVQK(SEQ ID NO:62)。在一些实施方案中,[N2]-[N3]为或包含YPAEVVQK(SEQ ID NO:943)、YPPSLVQK(SEQ ID NO:946)、NKAEVVQK(SEQ ID NO:947)、YTAEVVQK(SEQ ID NO:948)、YPAEVEQK(SEQ ID NO:949)、YPAEEVQK(SEQ ID NO:950)、YQAEVVQK(SEQ ID NO:951)、YTPSLVQK(SEQ ID NO:952)、YPPSLEQK(SEQ ID NO:953)、YPPSLVKK(SEQ ID NO:954)、

YPAEVVKK(SEQ ID NO:955)、YPAEVVHK(SEQ ID NO:956)、YPAAVVQK(SEQ ID NO:957)、NPAEVVQK(SEQ ID NO:958)、YPAEVVQQ(SEQ ID NO:959)、CPAEVVQK(SEQ ID NO:960)、YTAEVVKK(SEQ ID NO:961)、YPAEVLQK(SEQ ID NO:962)或YQAEEVQK(SEQ ID NO:963);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N2]-[N3]为YPAEVVQK(SEQ ID NO:943)。

[1177] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的AAV衣壳变体进一步包含[N1],其包含位置X_D、X_E和X_F。在一些实施方案中,[N1]的位置X_D为Q、T、S、A、I、L或H。在一些实施方案中,[N1]的位置X_E为S、G、A或R。在一些实施方案中,[N1]的位置X_F为S、K、L、R、A或T。在一些实施方案中,[N1]包含SK、SL、SS、SR、GA、GS、AS、ST、RS、QS、TS、AG、IG、QA、LG、HS、LS或QR。在一些实施方案中,[N1]为或包含QSS、QSK、TSL、SSS、QSR、AGA、IGS、QAS、ASS、LGS、QST、HSS、LSS或QRS。在一些实施方案中,[N1]的氨基酸序列为QSS。在一些实施方案中,[N1]-[N2]包含SSYPA(SEQ ID NO:63)、SKYPA(SEQ ID NO:64)、SLYPA(SEQ ID NO:65)、SRYPA(SEQ ID NO:66)、SSYPP(SEQ ID NO:67)、GAYPA(SEQ ID NO:68)、GSYPA(SEQ ID NO:69)、ASYPA(SEQ ID NO:70)、STNKA(SEQ ID NO:71)、SSYTA(SEQ ID NO:72)、SSYQA(SEQ ID NO:73)、SSYTP(SEQ ID NO:74)、SSNPA(SEQ ID NO:75)、SLCPA(SEQ ID NO:76)、RSYTA(SEQ ID NO:77)或SSTHA(SEQ ID NO:78)。在一些实施方案中,[N1]-[N2]包含SSYPAE(SEQ ID NO:79)、SKYPAE(SEQ ID NO:80)、SLYPAE(SEQ ID NO:81)、SRYPAE(SEQ ID NO:82)、SSYPPS(SEQ ID NO:83)、GAYPAE(SEQ ID NO:84)、GSYPAE(SEQ ID NO:85)、ASYPAE(SEQ ID NO:86)、STNKAE(SEQ ID NO:87)、SSYTAE(SEQ ID NO:88)、SSYQAE(SEQ ID NO:89)、SSYTPS(SEQ ID NO:90)、SSYPAA(SEQ ID NO:91)、SSNPAE(SEQ ID NO:92)、SLCPAE(SEQ ID NO:93)、RSYTAE(SEQ ID NO:94)、SSTHAS(SEQ ID NO:95)。在一些实施方案中,[N1]-[N2]为或包含QSSYPAEV(SEQ ID NO:96)、QSKYPAEV(SEQ ID NO:97)、TSLYPAEV(SEQ ID NO:98)、SSSYPAEV(SEQ ID NO:99)、QSRYPAEV(SEQ ID NO:100)、QSSYPPSL(SEQ ID NO:101)、AGAYPAEV(SEQ ID NO:102)、IGSYPAEV(SEQ ID NO:103)、QASYPAEV(SEQ ID NO:104)、ASSYPAEV(SEQ ID NO:105)、LGSYPAEV(SEQ ID NO:106)、QSTNKAEV(SEQ ID NO:107)、HSSYPAEV(SEQ ID NO:108)、SSSYTAEV(SEQ ID NO:109)、TSLYPAEE(SEQ ID NO:110)、ASSYQAEV(SEQ ID NO:111)、QSSYTPSL(SEQ ID NO:112)、QSRYPAEE(SEQ ID NO:113)、LSSYQAEV(SEQ ID NO:114)、HSSYPAAV(SEQ ID NO:115)、QSSNPAEV(SEQ ID NO:116)、QSSYTAEV(SEQ ID NO:117)、TSLCPAEV(SEQ ID NO:118)、QRSYTAEV(SEQ ID NO:119)或QSSYQAEE(SEQ ID NO:120);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]为QSSYPAEV(SEQ ID NO:96)。在一些实施方案中,[N1]-[N2]-[N3]包含SSYPAEVVQ(SEQ ID NO:121)、SKYPAEVVQ(SEQ ID NO:122)、SLYPAEVVQ(SEQ ID NO:123)、SRYPAEVVQ(SEQ ID NO:124)、SSYPPSLVQ(SEQ ID NO:125)、GAYPAEVVQ(SEQ ID NO:126)、GSYPAEVVQ(SEQ ID

NO:127)、ASYPAEVVQ(SEQ ID NO:128)、STNKAEVVQ(SEQ ID NO:129)、SSYTAEVVQ(SEQ ID NO:130)、SKYPAEVEQ(SEQ ID NO:131)、SLYPAEEVQ(SEQ ID NO:132)、SSYQAEVVQ(SEQ ID NO:133)、SSYTPSLVQ(SEQ ID NO:134)、SRYPAEEVQ(SEQ ID NO:135)、SSYPPSLEQ(SEQ ID NO:136)、SSYPPSLVK(SEQ ID NO:140)、SSYPAEVVK(SEQ ID NO:141)、SKYPAEVVH(SEQ ID NO:142)、SSYPAAVVQ(SEQ ID NO:143)、SSNPAEVVQ(SEQ ID NO:144)、SLCPAEVVQ(SEQ ID NO:145)、RSYTAEVVQ(SEQ ID NO:146)、SSYTAEVVK(SEQ ID NO:147)、SSYPAEVLQ(SEQ ID NO:148)或SSYQAEEVQ(SEQ ID NO:149)。在一些实施方案中,[N1]-[N2]-[N3]为或包含QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、QSSYPPSLVQK(SEQ ID NO:155)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)、QSTNKAEVVQK(SEQ ID NO:161)、HSSYPAEVVQK(SEQ ID NO:162)、SSSYTAEVVQK(SEQ ID NO:163)、QSKYPAEVEQK(SEQ ID NO:164)、TSLYPAEEVQK(SEQ ID NO:165)、ASSYQAEVVQK(SEQ ID NO:166)、QSSYTPSLVQK(SEQ ID NO:167)、QSRYPAEEVQK(SEQ ID NO:168)、QSSYPPSLEQK(SEQ ID NO:169)、QSSYPPSLVKK(SEQ ID NO:170)、LSSYQAEVVQK(SEQ ID NO:171)、SSSYPAEVVKK(SEQ ID NO:172)、QSKYPAEVVHK(SEQ ID NO:173)、HSSYPAAVVQK(SEQ ID NO:174)、QSSNPAEVVQK(SEQ ID NO:175)、SSSYPAEVVQQ(SEQ ID NO:176)、QSSYTAEVVQK(SEQ ID NO:177)、TSLCPAEVVQK(SEQ ID NO:178)、QRSYTAEVVQK(SEQ ID NO:179)、QSSYTAEVVKK(SEQ ID NO:180)、HSSYPAEVLQK(SEQ ID NO:181)或QSSYQAEEVQK(SEQ ID NO:182);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]-[N3]为QSSYPAEVVQK(SEQ ID NO:150)。

[1178] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的AAV衣壳变体进一步包含[N0],其中[N0]包含位置X_A、X_B和X_C。在一些实施方案中,[N0]的位置X_A为T、I或N。在一些实施方案中,[N0]的位置X_B为N。在一些实施方案中,[N0]的位置X_C为N、T、S或K。在一些实施方案中,[N0]包含TN、IN、NN、NT、NS或NK。在一些实施方案中,[N0]为或包含TNN、TNT、INN、TNS、NNN或TNK。在一些实施方案中,[N0]为TNN。在一些实施方案中,[N0]-[N1]为或包含TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNQST(SEQ ID NO:193)、TNNHSS(SEQ ID NO:194)、TNNQSK(SEQ ID NO:184)、TNNLSS(SEQ ID NO:195)、INNQSS(SEQ ID NO:196)、TNSQSS(SEQ ID NO:197)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)、TNNQRS(SEQ ID NO:200)或TNKQAS(SEQ ID NO:201);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]为TNNQSS(SEQ ID NO:

183)。在一些实施方案中,[NO]-[N1]-[N2]-[N3]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVVQK(SEQ ID NO:510)、TNNQSSYPPSLVQK(SEQ ID NO:512)、TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQK(SEQ ID NO:517)、TNTASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNQSTNKAEVVQK(SEQ ID NO:524)、TNNHSSYPAEVVQK(SEQ ID NO:525)、TNNSSSYTAEVVQK(SEQ ID NO:526)、TNNQSKYPAEVEQK(SEQ ID NO:529)、TNNTSLYPAEEVQK(SEQ ID NO:530)、TNTASSYQAEVVQK(SEQ ID NO:531)、TNNQSSYTPSLVQK(SEQ ID NO:533)、TNNQSRYPAEEVQK(SEQ ID NO:534)、TNNQSSYPPSLEQK(SEQ ID NO:535)、TNNQSSYPPSLVKK(SEQ ID NO:536)、TNNLSSYQAEVVQK(SEQ ID NO:539)、TNNSSSYPAEVVKK(SEQ ID NO:540)、TNNQSKYPAEVVHK(SEQ ID NO:542)、INNQSSYPAEVVQK(SEQ ID NO:543)、TNNHSSYPAAVVQK(SEQ ID NO:545)、TNSQSSNPAEVVQK(SEQ ID NO:548)、TNNSSSYPAEVVQQ(SEQ ID NO:551)、NNNQSRYPAEVVQK(SEQ ID NO:552)、TNNQSSYTAEVVQK(SEQ ID NO:553)、TNNTSLCPAEVVQK(SEQ ID NO:554)、TNSTSLYPAEVVQK(SEQ ID NO:556)、TNNQRSYTAEVVQK(SEQ ID NO:557)、TNNQSSYTAEVVKK(SEQ ID NO:558)、TNNHSSYPAEVLQK(SEQ ID NO:560)、TNNQSSYQAEEVQK(SEQ ID NO:562)或TNKQASYPAEVVQK(SEQ ID NO:563);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[NO]-[N1]-[N2]-[N3]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1179] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的AAV衣壳变体进一步包含[N4],其包含位置X_G和X_H。在一些实施方案中,[N4]的位置X_G为T、P或N。在一些实施方案中,[N4]的位置X_H为A。在一些实施方案中,[N4]为或包含TA、PA或NA。在一些实施方案中,[N4]为TA。在一些实施方案中,[N3]-[N4]为或包含VQKT A(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKKTA(SE Q ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SEQ ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571);包含任何上述氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N3]-[N4]为VQKTA(SE Q ID NO:564)。在一些实施方案中,[NO]-[N1]-[N2]-[N3]-[N4]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNNQSKYPAE VVQKTA(SEQ ID NO:1538)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEVVQKTA(SEQ ID NO:1539)、TNNQS RYPAEVVQKTA(SEQ ID NO:1327)、TNNQSSYPPSLVQKTA(SEQ ID NO:1300)、TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNIGSYPAEVVQKTA(SEQ ID NO:1112)、TNNQASYPAEVVQ KTA(SEQ ID NO:1194)、TNTASSYPAEVVQKTA(SEQ ID NO:1575)、TNNLGSYPAEVVQKTA(SEQ ID NO:1027)、TNNQSTNKAEVVQKTA(SEQ ID NO:1578)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TN NSSSYTAEVVQKTA(SEQ ID NO:1214)、TNNQSKYPAEVEQKT A(SEQ ID

NO:1254)、TNNTSLYPAEEVQKTA(SEQ ID NO:1583)、TNTASSYQAEVVQKTA(SEQ ID NO:1584)、TNNQSSYTPS LVQKTA(SEQ ID NO:1585)、TNNQSRYPAEEVQKTA(SEQ ID NO:1342)、TNNQSSYPPSLEQKTA(SEQ ID NO:1590)、TNNQS SYPPSLVKKTA(SEQ ID NO:1591)、TNNLSSYQAEVVQKTA(SEQ ID NO:1592)、TNNQSSYPPSLVQKPA(SEQ ID NO:1593)、TNNSSSYPAEVVKKTA(SEQ ID NO:1331)、TNNQSKYPAEVVH KTA(SEQ ID NO:1453)、TNNSSSYPAEVVQKPA(SEQ ID NO:1142)、INNQSSYPAEVVQKTA(SEQ ID NO:1024)、TNNHSSYPAAVVQKTA(SEQ ID NO:1598)、TNSQSSNPAEVVQKTA(SEQ ID NO:1599)、TNNSSSYPAEVVQQTA(SEQ ID NO:1419)、NN NQSRYPAEVVQKTA(SEQ ID NO:1601)、TNNQSSYTAEVVQKN A(SEQ ID NO:1602)、TNNTSLCPAEVVQKTA(SEQ ID NO:1603)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)、TNNQRSYTAE VVQKTA(SEQ ID NO:1604)、TNNQSSYTAEVVKKTA(SEQ ID NO:1606)、TNNHSSYPAEVLQKTA(SEQ ID NO:1607)、TNNQS SYQAEEVQKTA(SEQ ID NO:1608)或TNKQASYPAEVVQKTA(SEQ ID NO:1587);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[NO]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1180] 在一些实施方案中,本文所述的AAV衣壳变体包含式[N2]-[N3],其中[N2]包含位置X1、X2、X3、X4和X5且[N3]包含VQK或VQN的氨基酸序列。在一些实施方案中,[N3]包含氨基酸序列VQK。在一些实施方案中,[N3]为氨基酸序列VQK。在一些实施方案中,[N2]的位置X1为Y或T。在一些实施方案中,[N2]的位置X2为Q、T、P或E。在一些实施方案中,[N2]的位置X3为A。在一些实施方案中,[N2]的位置X4为E或D。在一些实施方案中,[N2]的位置X4为E或D。在一些实施方案中,[N2]的位置X5为V或E。在一些实施方案中,[N2]的位置X2为P、E、K、T或Q。在一些实施方案中,[N2]的位置X3为A或P。在一些实施方案中,[N2]的位置X4为E、S、D或A。在一些实施方案中,[N2]的位置X5为V、L或E。在一些实施方案中,[N2]包含位置X1处的Y。在一些实施方案中,[N2]包含位置X2处的P。在一些实施方案中,[N2]包含位置X3处的A。在一些实施方案中,[N2]包含位置X4处的E。在一些实施方案中,[N2]包含位置X5处的V。在一些实施方案中,[N2]包含YP、YQ、YT、TE、QA、TA、PA、EA、EV、EE、DV、AE或AD。在一些实施方案中,[N2]包含YPA、YQA、YTA、TEA、QAE、TAE、PAE、EAE、PAD、AEV、AEE或ADV。在一些实施方案中,[N2]包含YPAE(SEQ ID NO:21)、YQAE(SEQ ID NO:25)、YTAE(SEQ ID NO:24)、TEAE(SEQ ID NO:587)、YPAD(SEQ ID NO:588)、QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)、EAEV(SEQ ID NO:590)或PADV(SEQ ID NO:19)。在一些实施方案中,[N2]为或包含YPAEV(SEQ ID NO:1)、YQAEV(SEQ ID NO:6)、YTAEV(SEQ ID NO:4)、YPAEE(SEQ ID NO:5)、TEAEV(SEQ ID NO:12)或YPADV(SEQ ID NO:13)。在一些实施方案中,[N2]为YPAEV(SEQ ID NO:1)。在一些实施方案中,[N2]-[N3]包含AEVVQK(SEQ ID NO:36)、AEEVQK(SEQ ID NO:39)、AEVVQN(SEQ ID NO:591)或ADVVQK(SEQ ID NO:593)。在一些实施方案中,[N2]-[N3]包含PAEVVQN(SEQ ID NO:594)、QAEVVQK(SEQ ID NO:52)、TAEVVQK(SEQ ID NO:49)、PAEVVQK(SEQ ID NO:20)、PAEEVQK(SEQ ID NO:51)、EAEVVQK(SEQ ID NO:595)或PADVVQK(SEQ ID NO:596)。在一些实施方案中,[N2]-[N3]为或包含YPAEVVQK(SEQ ID

NO:943)、YQAEVVQK(SEQ ID NO:951)、YTAEVVQK(SEQ ID NO:948)、YPAEEVQK(SEQ ID NO:950)、YPAEVVQN(SEQ ID NO:964)、TEAEVVQK(SEQ ID NO:965)或YPADVVQK(SEQ ID NO:966);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N2]-[N3]为YPAEVVQK(SEQ ID NO:943)。

[1181] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的AAV衣壳变体进一步包含[N1],其包含位置X_D、X_E和X_F。在一些实施方案中,[N1]的位置X_D为Q或S。在一些实施方案中,[N1]的位置X_E为S、L、A或T。在一些实施方案中,[N1]的位置X_F为S、Y或T。在一些实施方案中,[N1]包含QS、SL、SA、QT、LS、LY、AT、TS或SS。在一些实施方案中,[N1]为或包含QSS、SLS、SLY、SAT或QTS。在一些实施方案中,[N1]为QSS。在一些实施方案中,[N1]-[N2]包含SSYPA(SEQ ID NO:63)、LSYQA(SEQ ID NO:597)、LSYTA(SEQ ID NO:598)、LYYPA(SEQ ID NO:600)、ATYPA(SEQ ID NO:601)、LSYPA(SEQ ID NO:603)或TSTEA(SEQ ID NO:605)。在一些实施方案中,[N1]-[N2]包含SSYPAE(SEQ ID NO:79)、LSYQAE(SEQ ID NO:607)、LSYTAE(SEQ ID NO:610)、LYYPAE(SEQ ID NO:611)、ATYPAE(SEQ ID NO:613)、LSYPAE(SEQ ID NO:616)、TSTEAE(SEQ ID NO:619)或LSYPAD(SEQ ID NO:621)。在一些实施方案中,[N1]-[N2]为或包含QSSYPAEV(SEQ ID NO:96)、SLSYQAEV(SEQ ID NO:622)、SLSYTAEV(SEQ ID NO:623)、SLYYPAEV(SEQ ID NO:624)、SATYPAEV(SEQ ID NO:625)、SLSYPAEV(SEQ ID NO:629)、SLSYPAEE(SEQ ID NO:632)、QTSTEAEV(SEQ ID NO:633)或SLSYPADV(SEQ ID NO:634);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6或7个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]为QSSYPAEV(SEQ ID NO:96)。在一些实施方案中,[N1]-[N2]-[N3]为或包含QSSYPAEVVQK(SEQ ID NO:150)、SLSYQAEVVQK(SEQ ID NO:635)、SLSYTAEVVQK(SEQ ID NO:637)、SLYYPAEVVQK(SEQ ID NO:639)、SATYPAEVVQK(SEQ ID NO:641)、SLSYPAEVVQK(SEQ ID NO:642)、SLSYPAEEVQK(SEQ ID NO:643)、SLSYPAEVVQN(SEQ ID NO:644)、QTSTEAEVVQK(SEQ ID NO:645)或SLSYPADVVQK(SEQ ID NO:646);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N1]-[N2]-[N3]为QSSYPAEVVQK(SEQ ID NO:150)。

[1182] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的AAV衣壳变体进一步包含[N0],其中[N0]包含位置X_A、X_B和X_C。在一些实施方案中,[N0]的位置X_A为T。在一些实施方案中,[N0]的位置X_B为N。在一些实施方案中,[N0]的位置X_C为N、T、S或K。在一些实施方案中,[N0]包含TN、NS、NT、NN或NK。在一些实施方案中,[N0]为或包含TNS、TNT、TNN或TNK。在一些实施方案中,[N0]为TNN。在一些实施方案中,[N0]-[N1]为或包含TNNQSS(SEQ ID NO:183)、TNSSLS(SEQ ID NO:647)、TNSSLY(SEQ ID NO:648)、TNTSAT(SEQ ID NO:649)、TNNQTS(SEQ

ID NO:650)或TNKSAT(SEQ ID NO:651);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[N0]-[N1]-[N2]-[N3]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNSSLSYQAEVVQK(SEQ ID NO:652)、TNSSLSYTAEVVQK(SEQ ID NO:654)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNTSATYPAEVVQK(SEQ ID NO:656)、TNSSLSYPAEVVQK(SEQ ID NO:657)、TNSSLSYPAEEVQK(SEQ ID NO:658)、TNSSLSYPAEVVQN(SEQ ID NO:660)、TNNQTSTEAEVVQK(SEQ ID NO:662)、TNKSATYPAEVVQK(SEQ ID NO:663)或TNSSLSYPADVVQK(SEQ ID NO:665);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]-[N2]-[N3]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1183] 在一些实施方案中,包括包含式[N2]-[N3]的氨基酸序列的AAV衣壳变体进一步包含[N4],其包含位置X_G和X_H。在一些实施方案中,[N4]的位置X_G为T、P或N。在一些实施方案中,[N4]的位置X_H为A或D。在一些实施方案中,[N4]为或包含TA、TD、PA或NA。在一些实施方案中,[N4]为TA。在一些实施方案中,[N3]-[N4]为或包含VQKTA(SEQ ID NO:564)、EQKTA(SEQ ID NO:565)、VKK TA(SEQ ID NO:566)、VQKPA(SEQ ID NO:567)、VHKTA(SE Q ID NO:568)、VQQTA(SEQ ID NO:569)、VQKNA(SEQ ID NO:570)或LQKTA(SEQ ID NO:571);包含任何上述氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含至少一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N3]-[N4]为VQKTA(SEQ ID NO:564)。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNSSLSYQAEVVQKTA(SEQ ID NO:2064)、TNSSLSYTAEVVQ KTA(SEQ ID NO:2065)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNTSATYPAEVVQKTA(SEQ ID NO:2067)、TNSSLSYPA EVVQKTA(SEQ ID NO:2068)、TNSSLSYPAEEVQKTA(SEQ ID NO:2069)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQNTA(SEQ ID NO:2071)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)、TNNQTSTEAEVVQKTA(SEQ ID NO:2074)、TNKSATYPAEVVQ KTA(SEQ ID NO:2075)或TNSSLSYPADVVQKTA(SEQ ID NO:2076);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKT(SEQ ID NO:737)。

[1184] 在一些实施方案中,[N2]-[N3]存在于AAV衣壳变体的环VIII中。在一些实施方案

中,[N0]、[N1]和/或[N4]存在于AAV衣壳变体的环VIII中。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]存在于AAV衣壳变体的环VIII中。在一些实施方案中,环VIII包含位置571-592,这些位置根据SEQ ID NO:138编号;或位置571-599,这些位置根据SEQ ID NO:982编号。

[1185] 在一些实施方案中,[N0]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在一些实施方案中,[N0]置换位置571-573(例如,T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N0]紧接在位置570之后存在,且[N0]置换位置571-573(例如,氨基酸T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N1]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在一些实施方案中,[N1]置换位置574-576(例如,Q574、S575和S576),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N1]紧接在位置573之后存在,且[N1]置换位置574-576(例如,Q574、S575和S576),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N2]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在一些实施方案中,[N2]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在一些实施方案中,[N2]紧接在位置576之后存在,且[N2]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N2]-[N3]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在一些实施方案中,[N2]-[N3]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在一些实施方案中,[N2]-[N3]紧接在位置576之后存在,且[N2]-[N3]置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N2]-[N3]-[N4]紧接在相对于SEQ ID NO:138编号的位置576之后存在。在一些实施方案中,[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N2]-[N3]-[N4]紧接在位置576之后存在,且[N2]-[N3]-[N4]置换位置577-579(例如,T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N1]-[N2]-[N3]-[N4]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在一些实施方案中,[N1]-[N2]-[N3]-[N4]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N1]-[N2]-[N3]-[N4]紧接在位置573之后存在,且[N1]-[N2]-[N3]-[N4]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]紧接在位置570之后存在。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]紧接在位置570之后存在,且[N0]-[N1]-[N2]-[N3]-[N4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中[N2]-[N3]置换相对于SEQ ID NO:138编号的位置577(例如,T577)。

[1186] 在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,AAV衣壳变体包含位置577处非T的氨基酸。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,AAV衣壳变体包含位置577处的Y。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[N2]的X1存在于位置577(例如,T577)处,且[N2]的位置X2和X3紧接在位置577之

后存在。在一些实施方案中,[N3]紧接在[N2]之后存在。

[1187] 在一些实施方案中,[N0]的X_A存在于位置571处,[N0]的X_B存在于位置572处,且[N0]的X_C存在于位置573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N1]的X_D存在于位置574处,[N1]的X_E存在于位置575处,且[N1]的X_F存在于位置576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]的X1存在于位置577处,[N2]的X2存在于位置578处,[N2]的X3存在于位置579处,[N2]的X4存在于位置580处,且[N2]的X5存在于位置581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N4]的X_G存在于位置585处且[N4]的X_H存在于位置586处,这些位置根据SEQ ID NO:982编号。

[1188] 在一些实施方案中,[N0]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N1]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]存在于位置577-581处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N3]存在于位置582-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N4]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[1189] 在一些实施方案中,[N1]紧接在[N0]之后存在。在一些实施方案中,[N2]紧接在[N1]之后存在。在一些实施方案中,[N3]紧接在[N2]之后存在。在一些实施方案中,[N4]紧接在[N3]之后存在。

[1190] 在一些实施方案中,AAV衣壳变体自N端至C端包含[N2]-[N3]。在一些实施方案中,AAV衣壳变体自N端至C端包含[N1]-[N2]-[N3]。在一些实施方案中,AAV衣壳变体自N端至C端包含[N1]-[N2]-[N3]-[N4]。在一些实施方案中,AAV衣壳变体自N端至C端包含[N0]-[N1]-[N2]-[N3]。在一些实施方案中,AAV衣壳变体自N端至C端包含[N0]-[N1]-[N2]-[N3]-[N4]。

[1191] 在一些实施方案中,[N2]-[N3]为YPAEVVQK(SEQ ID NO:943),其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577。在一些实施方案中,[N0]-[N1]-[N2]-[N3]-[N4]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)且紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中[N2]-[N3](YPAEVVQK(SEQ ID NO:943))置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。在一些实施方案中,[N2]-[N3]为YPAEVVQK,其中[N2]-[N3]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。

[1192] 在一些实施方案中,本文所述的AAV衣壳变体包含具有式[B]-[C]的氨基酸序列,其中[B]包含位置X1、X2和X3,且[C]包含氨基酸序列YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,[B]的位置X1为Q、T、S、A、I、L或H。在一些实施方案中,[B]的位置X1为Q、T、S、A或H。在一些实施方案中,[B]的位置X2为S、G或A。在一些实施方案中,[B]的位置X2为S或G。在一些实施方案中,[B]的位置X3为S、K、L、R或A。在一些实施方案中,[B]的位置X3为S、K、L或R。在一些实施方案中,[B]包含位置X1处的Q。在一些实施方案中,[B]包含位置X2处的S。在一些实施方案中,[B]包含位置X3处的S。在一些实施方案中,[B]包含QS、TS、SS、AG、IG、QA、AS、LG、HS、SK、SL、SR、GA或GS。在一些实施方案中,[B]为或包含QSS、TSL、SSS、QSR、QSK、AGA、IGS、QAS、ASS、LGS或HSS。在一些实施方案中,[B]为QSS。在一些实施方案中,[B]-[C]包含

SSYPAEVVQK(SEQ ID NO:572)、SKYPAEVVQK(SEQ ID NO:573)、SLYPAEVVQK(SEQ ID NO:574)、SRYPAEVVQK(SEQ ID NO:575)、GAYPAEVVQK(SEQ ID NO:576)、GSYPAEVVQK(SEQ ID NO:580)或ASYPAEVVQK(SEQ ID NO:582)。在一些实施方案中,[B]-[C]为或包含QSSYPAEVVQK(SEQ ID NO:150)、QSKYPAEVVQK(SEQ ID NO:151)、TSLYPAEVVQK(SEQ ID NO:152)、SSSYPAEVVQK(SEQ ID NO:153)、QSRYPAEVVQK(SEQ ID NO:154)、AGAYPAEVVQK(SEQ ID NO:156)、IGSYPAEVVQK(SEQ ID NO:157)、QASYPAEVVQK(SEQ ID NO:158)、ASSYPAEVVQK(SEQ ID NO:159)、LGSYPAEVVQK(SEQ ID NO:160)或HSSYPAEVVQK(SEQ ID NO:162);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[B]-[C]为QSSYPAEVVQK(SEQ ID NO:150)。

[1193] 在一些实施方案中,包含式[B]-[C]的AAV衣壳变体进一步包含[A],其包含位置X_A、X_B和X_C。在一些实施方案中,[A]的位置X_A为T、I或N。在一些实施方案中,[A]的位置X_B为N。在一些实施方案中,[A]的位置X_C为N、T、S或K。在一些实施方案中,[A]包含TN、IN、NN、NT、NS或NK。在一些实施方案中,[A]为或包含TNN、TNT、INN、NNN、TNS或TNK。在一些实施方案中,[A]为TNN。在一些实施方案中,[A]-[B]为或包含TNNQSS(SEQ ID NO:183)、TNNQSK(SEQ ID NO:184)、TNNTSL(SEQ ID NO:185)、TNNSSS(SEQ ID NO:186)、TNNQSR(SEQ ID NO:187)、TNNAGA(SEQ ID NO:188)、TNNIGS(SEQ ID NO:189)、TNNQAS(SEQ ID NO:190)、TNTASS(SEQ ID NO:191)、TNNLGS(SEQ ID NO:192)、TNNHSS(SEQ ID NO:194)、INNQSS(SEQ ID NO:196)、NNNQSR(SEQ ID NO:198)、TNSTSL(SEQ ID NO:199)或TNKQAS(SEQ ID NO:201);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[A]-[B]-[C]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNNQSKYPAEVVQK(SEQ ID NO:503)、TNNTSLYPAEVVQK(SEQ ID NO:506)、TNNSSSYPAEVVQK(SEQ ID NO:508)、TNNQSRYPAEVVQK(SEQ ID NO:510)、TNNAGAYPAEVVQK(SEQ ID NO:513)、TNNIGSYPAEVVQK(SEQ ID NO:514)、TNNQASYPAEVVQK(SEQ ID NO:517)、TNTASSYPAEVVQK(SEQ ID NO:520)、TNNLGSYPAEVVQK(SEQ ID NO:523)、TNNHSSYPAEVVQK(SEQ ID NO:525)、INNQSSYPAEVVQK(SEQ ID NO:543)、NNNQSRYPAEVVQK(SEQ ID NO:552)、TNSTSLYPAEVVQK(SEQ ID NO:556)或TNKQASYPAEVVQK(SEQ ID NO:563);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]-[C]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1194] 在一些实施方案中,包含式[B]-[C]的AAV衣壳变体进一步包含[D],其中[D]包含位置X4和X5。在一些实施方案中,[D]的位置X4为T或N。在一些实施方案中,[D]的位置X5为

A。在一些实施方案中,[D]为或包含TA或PA。在一些实施方案中,[D]为TA。在一些实施方案中,[C]-[D]为或包含YPAEVVQKTA(SEQ ID NO:584)或YPAEVVQKPA(SEQ ID NO:586);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[C]-[D]为YPAEVVQKTA(SEQ ID NO:584)。在一些实施方案中,[A]-[B]-[C]-[D]为或包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)、TNNQSKYPAEVVQKTA(SEQ ID NO:1538)、TNNTSLYPAEVVQKTA(SEQ ID NO:1232)、TNNSSSYPAEVVQKTA(SEQID NO:1539)、TNNQSRYPAEVVQKTA(SEQ ID NO:1327)、TNNAGAYPAEVVQKTA(SEQ ID NO:1021)、TNNIGSYPAEVVQK TA(SEQ ID NO:1112)、TNNQASYPAEVVQKTA(SEQ ID NO:1194)、TNTASSYPAEVVQKTA(SEQ ID NO:1575)、TNNLGSYP AEVVQKTA(SEQ ID NO:1027)、TNNHSSYPAEVVQKTA(SEQ ID NO:1310)、TNNSSSYPAEVVQKPA(SEQ ID NO:1142)、INN QSSYPAEVVQKTA(SEQ ID NO:1024)、NNNQSRYPAEVVQKTA(SEQ ID NO:1601)、TNSTSLYPAEVVQKTA(SEQ ID NO:1605)或TNKQASYPAEVVQKTA(SEQ ID NO:1587);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]-[C]-[D]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1195] 在一些实施方案中,本文所述的AAV衣壳变体包含具有式[B]-[C]的氨基酸序列,其中[B]包含位置X1、X2和X3,且[C]包含氨基酸序列YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,[B]的位置X1为Q或S。在一些实施方案中,[B]的位置X2为S、L或A。在一些实施方案中,[B]的位置X3为S、Y或T。在一些实施方案中,[B]包含位置X1处的Q。在一些实施方案中,[B]包含位置X2处的S。在一些实施方案中,[B]包含位置X3处的S。在一些实施方案中,[B]包含QS、SL、SA、LY、AT、LS或SS。在一些实施方案中,[B]为或包含QSS、SLY、SAT或SLS。在一些实施方案中,[B]为QSS。在一些实施方案中,[B]-[C]包含SSYPAEVVQK(SEQ ID NO:572)、LYYPAEVVQK(SEQ ID NO:702)、ATYPAEVVQK(SEQ ID NO:718)或LSYPAEVVQK(SEQ ID NO:703)。在一些中,[B]-[C]为或包含QSSYPAEVVQK(SEQ ID NO:150)、SLYYPAEVVQK(SEQ ID NO:639)、SATYPAEVVQK(SEQ ID NO:641)或SLSYPAEVVQK(SEQ ID NO:642);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些中,[B]-[C]为QSSYPAEVVQK(SEQ ID NO:150)。

[1196] 在一些实施方案中,包含式[B]-[C]的AAV衣壳变体进一步包含[A],其包含位置X_A、X_B和X_C。在一些实施方案中,[A]的位置X_A为T。在一些实施方案中,[A]的位置X_B为N。在一些实施方案中,[A]的位置X_C为N、T、S或K。在一些实施方案中,[A]包含TN、NS、NT、NK或NN。在一些实施方案中,[A]为或包含TNN、TNS、TNT或TNK。在一些实施方案中,[A]为TNN。在一些实

施方案中,[A]-[B]为或包含TNNQSS(SEQ ID NO:183)、TNSSLY(SEQ ID NO:648)、TNTSAT(SEQ ID NO:649)、TNSSLS(SEQ ID NO:647)或TNKSAT(SEQ ID NO:651);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4或5个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]为TNNQSS(SEQ ID NO:183)。在一些实施方案中,[A]-[B]-[C]为或包含TNNQSSYPAEVVQK(SEQ ID NO:500)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNTSATYPAEVVQK(SEQ ID NO:656)、TNSSLSYPAEVVQK(SEQ ID NO:657)或TNKSATYPAEVVQK(SEQ ID NO:663);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]-[C]为TNNQSSYPAEVVQK(SEQ ID NO:500)。

[1197] 在一些实施方案中,包含式[B]-[C]的AAV衣壳变体进一步包含[D],其中[D]包含位置X4和X5。在一些实施方案中,[D]的位置X4为T、N或P。在一些实施方案中,[D]的位置X5为A或D。在一些实施方案中,[D]为或包含TA、TD、NA或PA。在一些实施方案中,[D]为TA。在一些实施方案中,[C]-[D]为或包含YPAEVVQKTA(SEQ ID NO:584)、YPAEVVQKTD(SEQ ID NO:719)、YPAEVVQ KNA(SEQ ID NO:724)或YPAEVVQKPA(SEQ ID NO:586);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[C]-[D]为YPAEVVQKTA(SEQ ID NO:584)。在一些实施方案中,[A]-[B]-[C]-[D]为或包含TNNQSSYPAEVV QKTA(SEQ ID NO:1533)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNTSATYPAEVVQKTA(SEQ ID NO:2067)、TNSSLSYP AEVVQKTA(SEQ ID NO:2068)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNS SLSYPAEVVQKPA(SEQ ID NO:2073)或TNKSATYPAEVVQKT A(SEQ ID NO:2075);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[A]-[B]-[C]-[D]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)。

[1198] 在一些实施方案中,[B]-[C]存在于AAV衣壳变体的环VIII中。在一些实施方案中,[A]和/或[D]存在于AAV衣壳变体的环VIII中。在一些实施方案中,[A]-[B]-[C]-[D]存在于AAV衣壳变体的环VIII中。在一些实施方案中,环VIII包含位置571-592,这些位置根据SEQ ID NO:138编号。在一些实施方案中,环VIII包含位置571-599,这些位置根据SEQ ID NO:982编号。

[1199] 在一些实施方案中,[A]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在

这些实施方案的任一个中,[A]置换位置571-573(例如,T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[A]紧接在位置570之后存在,且[A]置换位置571-573(例如,T571、N572和N573),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]置换位置574-576(例如,Q574、S575和S576)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]紧接在位置573之后存在,且[B]置换位置574-576(例如,Q574、S575和S576)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]紧接在位置576之后存在,其中[C]置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[B]-[C]紧接在位置573之后存在,且[B]-[C]置换位置574-577(例如,Q574、S575、S576和T577)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]-[D]紧接在位置576之后存在。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]-[D]置换位置577-579(例如,T577、T578和A579)。在一些实施方案中,相对于根据SEQ ID NO:138编号的参考序列,[C]-[D]紧接在位置576之后存在,且[C]-[D]置换位置577-579(例如,T577、T578和A579)。在一些实施方案中,[B]-[C]-[D]紧接在相对于SEQ ID NO:138编号的位置573之后存在。在一些实施方案中,[B]-[C]-[D]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[B]-[C]-[D]紧接在位置573之后存在,且[B]-[C]-[D]置换位置574-579(例如,Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[A]-[B]-[C]-[D]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在一些实施方案中,[A]-[B]-[C]-[D]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[A]-[B]-[C]-[D]紧接在位置570之后存在,且[A]-[B]-[C]-[D]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。

[1200] 在一些实施方案中,[A]的X_A存在于位置571处,[A]的X_B存在于位置572处,且[A]的X_C存在于位置573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]的X1存在于位置574处,[B]的X2存在于位置575处,且[B]的X3存在于位置576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[D]的X4存在于位置585处且[D]的位置X5存在于位置586处,这些位置根据SEQ ID NO:982编号。

[1201] 在一些实施方案中,[A]存在于位置571-573处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[B]存在于位置574-576处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,[D]存在于位置585-586处,这些位置根据SEQ ID NO:982编号。在一些实施方案中,

[A]-[B]-[C]-[D]存在于位置571-586处,这些位置根据SEQ ID NO:982编号。

[1202] 在一些实施方案中,[B]紧接在[A]之后存在。在一些实施方案中,[C]紧接在[B]之后存在。在一些实施方案中,[D]紧接在[C]之后存在。在一些实施方案中,AAV衣壳变体自N端至C端包含[B]-[C]。在一些实施方案中,AAV衣壳变体自N端至C端包含[B]-[C]。在一些实施方案中,AAV衣壳变体自N端至C端包含[A]-[B]-[C]。在一些实施方案中,AAV衣壳变体自N端至C端包含[B]-[C]-[D]。在一些实施方案中,AAV衣壳变体自N端至C端包含[A]-[B]-[C]-[D]。

[1203] 在一些实施方案中,[C]为YPAEVVQK(SEQ ID NO:943),其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在一些实施方案中,[A]-[B]-[C]-[D]为TNNQSSYPAEVVQKTA(SEQ ID NO:1533)且紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中[C](YPAEVVQK(SEQ ID NO:943))置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。在一些实施方案中,[C]为YPAEVVQK,其中[C]存在于位置577-584处,这些位置根据SEQ ID NO:982编号。

[1204] 在一些实施方案中,本文所述的AAV衣壳变体包含式[K1]-[K2],其中[K1]包含LSY或LYY,且[K2]包含位置X1、X2、X3和X4。在一些实施方案中,[K1]包含LSY。在一些实施方案中,[K2]的位置X1为Q、T或P。在一些实施方案中,[K2]的位置X2为A,在一些实施方案中,[K2]的位置X3为E或D。在一些实施方案中,[K2]的位置X4为V或E。在一些实施方案中,[K2]包含QA、TA、PA、EV、EE、DV、AE或AD。在一些实施方案中,[K2]包含QAE、TAE、PAE、PAD、AEV、AEE或ADV。在一些实施方案中,[K2]为或包含QAEV(SEQ ID NO:15)、TAEV(SEQ ID NO:16)、PAEV(SEQ ID NO:17)、PAEE(SEQ ID NO:18)或PADV(SEQ ID NO:19)。在一些实施方案中,[K1]-[K2]包含LSYQA(SEQ ID NO:597)、LSYTA(SEQ ID NO:598)、LYYPA(SEQ ID NO:600)或LSYPA(SEQ ID NO:603)。在一些实施方案中,[K1]-[K2]包含LSYQAE(SEQ ID NO:607)、LSYTAE(SEQ ID NO:610)、LYYPAE(SEQ ID NO:611)、LSYPAE(SEQ ID NO:616)或LSYPAD(SEQ ID NO:621)。在一些实施方案中,[K1]-[K2]为或包含LSYQAEV(SEQ ID NO:667)、LSYTAEV(SEQ ID NO:668)、LYYPAEV(SEQ ID NO:669)、LSYPAEV(SEQ ID NO:671)、LSYPAEE(SEQ ID NO:673)或LSYPADV(SEQ ID NO:674);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1205] 在一些实施方案中,包括包含式[K1]-[K2]的氨基酸序列的AAV衣壳变体进一步包含[K0],其包含TNNS(SEQ ID NO:14)。在一些实施方案中,[K0]-[K1]包含TNSSLS(SEQ ID NO:647)或TNSSLY(SEQ ID NO:648)。在一些实施方案中,[K0]-[K1]为或包含TNSSLSY(SEQ ID NO:676)或TNSSLYY(SEQ ID NO:678);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K0]-[K1]-[K2]包含TNSSLSYQA(SEQ ID NO:679)、TNSSLSYTA(SEQ ID NO:681)、TNSSLYYPA(SEQ ID NO:682)或TNSSLSYPA(SEQ ID NO:683)。

在一些实施方案中,[K0]-[K1]-[K2]包含TNSSLSYQAE(SEQ ID NO:684)、TNSSLSYTAE(SEQ ID NO:685)、TNSSLYYPAE(SEQ ID NO:686)、TNSSLSYPAE(SEQ ID NO:687)或TNSSLSYPAD(SEQ ID NO:689)。在一些实施方案中,[K0]-[K1]-[K2]为或包含TNSSLSYQAEV(SEQ ID NO:692)、TNSSLSYTAEV(SEQ ID NO:693)、TNSSLYYPAEV(SEQ ID NO:696)、TNSSLSYPAEV(SEQ ID NO:697)、TNSSLSYPAEE(SEQ ID NO:698)或TNSSLSYPADV(SEQ ID NO:699);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9或10个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1206] 在一些实施方案中,包括包含式[K1]-[K2]的氨基酸序列的AAV衣壳变体进一步包含[K3],其中[K3]包含位置X_A、X_B和X_C。在一些实施方案中,[K3]的位置X_A为V。在一些实施方案中,[K3]的位置X_B为Q。在一些实施方案中,[K3]的位置X_C为K或N。在一些实施方案中,[K3]包含VQ、QK或QN。在一些实施方案中,[K3]为或包含VQK或VQN。在一些实施方案中,K2]-[K3]为或包含QAEVVQK(SEQ ID NO:52)、TAEVVQK(SEQ ID NO:49)、PAEVVQK(SEQ ID NO:20)、PAEEVQK(SEQ ID NO:51)、PAEVVQN(SEQ ID NO:594)或PADVVQK(SEQ ID NO:596);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5或6个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K1]-[K2]-[K3]为或包含LSYQAEVVQK(SEQ ID NO:700)、LSYTAEVVQK(SEQ ID NO:701)、LYYPAEVVQK(SEQ ID NO:702)、LSYPAEVVQK(SEQ ID NO:703)、LSYPAEEVQK(SEQ ID NO:704)、LSYPAEVVQN(SEQ ID NO:706)或LSYPADVVQK(SEQ ID NO:708);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8或9个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K0]-[K1]-[K2]-[K3]为或包含TNSSLSYQAEVVQK(SEQ ID NO:652)、TNSSLSYTAEVVQK(SEQ ID NO:654)、TNSSLYYPAEVVQK(SEQ ID NO:655)、TNSSLSYPAEVVQK(SEQ ID NO:657)、TNSSLSYPAEEVQK(SEQ ID NO:658)、TNSSLSYPAEVVQN(SEQ ID NO:660)或TNSSLSYPADVVQK(SEQ ID NO:665);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12或13个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1207] 在一些实施方案中,包括包含式[K1]-[K2]的氨基酸序列的AAV衣壳变体进一步包含[K4],其中[K4]包含位置X_D和X_E。在一些实施方案中,[K4]的位置X_D为T、P或N。在一些实施方案中,[K4]的位置X_E为A或D。在一些实施方案中,[K4]为或包含TA、TD、PA或NA。在一些实施方案中,[K3]-[K4]为或包含VQKTA(SEQ ID NO:564)、VQKTD(SEQ ID NO:714)、VQNTA(SEQ ID NO:715)、VQKNA(SEQ ID NO:570)或VQKPA(SEQ ID NO:567);包含任何上述氨基酸序列的任何部分(例如,任何2、3或4个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述

氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,[K0]-[K1]-[K2]-[K3]-[K4]为或包含TNSSLSYQAEVVQKTA(SE Q ID NO:2064)、TNSSLSYTAEVVQKTA(SEQ ID NO:2065)、TNSSLYYPAEVVQKTA(SEQ ID NO:2066)、TNSSLSYPAEVVQK TA(SEQ ID NO:2068)、TNSSLSYPAEEVQKTA(SEQ ID NO:2069)、TNSSLSYPAEVVQKTD(SEQ ID NO:2070)、TNSSLSYPAEVVQNTA(SEQ ID NO:2071)、TNSSLSYPAEVVQKNA(SEQ ID NO:2072)、TNSSLSYPAEVVQKPA(SEQ ID NO:2073)或TNSSL SYPADVVQKTA(SEQ ID NO:2076);包含任何上述氨基酸序列的任何部分(例如,任何2、3、4、5、6、7、8、9、10、11、12、13、14或15个氨基酸,例如连续氨基酸)的氨基酸序列;相对于任何上述氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列;或相对于任一上述氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1208] 在一些实施方案中,[K1]-[K2]存在于AAV衣壳变体的环VIII中。在一些实施方案中,[K0]、[K3]和/或[K4]存在于AAV衣壳变体的环VIII中。在一些实施方案中,[K0]-[K1]-[K2]-[K3]-[K4]存在于AAV衣壳变体的环VIII中。在一些实施方案中,环VIII包含位置571-592,这些位置根据SEQ ID NO:138编号。在一些实施方案中,环VIII包含位置571-599,这些位置根据SEQ ID NO:982编号。

[1209] 在一些实施方案中,[K0]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在一些实施方案中,[K0]置换位置571-574(例如,T571、N572、N573和Q574),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K0]紧接在位置570之后存在,且[K0]置换位置571-574(例如,T571、N572、N573和Q574),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K1]紧接在相对于SEQ ID NO:138编号的位置574之后存在。在一些实施方案中,[K1]置换位置575-577(例如,S575、S576和T577),序列相对于SEQ ID NO:138。在一些实施方案中,[K1]紧接在位置574之后存在,其中[K1]置换位置575-577(例如,S575、S576和T577),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K1]-[K2]-[K3]紧接在相对于SEQ ID NO:138编号的位置574之后存在。在一些实施方案中,[K1]-[K2]-[K3]置换位置575-577(例如,S575、S576和T577),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K1]-[K2]-[K3]紧接在位置574之后存在,其中[K1]-[K2]-[K3]置换位置575-577(例如,S575、S576和T577),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K1]-[K2]-[K3]-[K4]紧接在相对于SEQ ID NO:138编号的位置574之后存在。在一些实施方案中,[K1]-[K2]-[K3]-[K4]置换位置575-579(例如,S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K1]-[K2]-[K3]-[K4]紧接在位置574之后存在,其中[K1]-[K2]-[K3]-[K4]置换位置575-579(例如,S575、S576、T577、T578和A579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K0]-[K1]-[K2]-[K3]-[K4]紧接在相对于SEQ ID NO:138编号的位置570之后存在。在一些实施方案中,[K0]-[K1]-[K2]-[K3]-[K4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和T579),这些位置相对于SEQ ID NO:138编号。在一些实施方案中,[K0]-[K1]-[K2]-[K3]-[K4]紧接在位置570之后存在,且[K0]-[K1]-[K2]-[K3]-[K4]置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和T579),这些位置相对于SEQ ID NO:138编号。

[1210] 在一些实施方案中,AAV衣壳变体自N端至C端包含[K1]-[K2]。在一些实施方案中,AAV衣壳变体自N端至C端包含[K0]-[K1]-[K2]。在一些实施方案中,AAV衣壳变体自N端至C端包含[K1]-[K2]-[K3]。在一些实施方案中,AAV衣壳变体自N端至C端包含[K0]-[K1]-[K2]-[K3]。在一些实施方案中,AAV衣壳变体自N端至C端包含[K1]-[K2]-[K3]-[K4]。在一些实施方案中,AAV衣壳变体自N端至C端包含[K0]-[K1]-[K2]-[K3]-[K4]。

[1211] 在一些实施方案中,本文所述的AAV衣壳变体包含氨基酸序列,所述氨基酸序列包含来自表1、2A、2B、9和15-20中所提供的任一序列的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸。在一些实施方案中,AAV衣壳变体包含氨基酸序列,所述氨基酸序列包含来自SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸。在一些实施方案中,AAV衣壳变体包含氨基酸序列,所述氨基酸序列包含来自SEQ ID NO:943或2064-2080中任一个的至少3、4、5、6、7、8、9、10、11、12、13、14或15个连续氨基酸。在一些实施方案中,AAV衣壳变体包含来自SEQ ID NO:943或946-966中任一个的至少3、4、5、6或7个连续氨基酸。在一些实施方案中,所述氨基酸序列存在于环VIII中。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置570、571、572、573、574、575或576之后存在。

[1212] 在一些实施方案中,3个连续氨基酸包含YPA。在一些实施方案中,4个连续氨基酸包含YPAE(SEQ ID NO:21)。在一些实施方案中,5个连续氨基酸包含YPAEV(SEQ ID NO:1)。在一些实施方案中,6个连续氨基酸包含YPAEVV(SEQ ID NO:725)。在一些实施方案中,7个连续氨基酸包含YPAEVVQ(SEQ ID NO:726)。在一些实施方案中,所述氨基酸序列包含YPAEVVQK(SEQ ID NO:943)。在一些实施方案中,所述氨基酸序列由YPAEVVQK(SEQ ID NO:943)组成。

[1213] 在一些实施方案中,本文所述的AAV衣壳变体包含氨基酸序列,所述氨基酸序列相对于表1、2A、2B、9和15-20中所提供的任一序列的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)。在一些实施方案中,AAV衣壳变体包含相对于表1、2A、2B、9和15-20中所提供的任一序列的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:943或2064-2080中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基

酸序列。在一些实施方案中,所述氨基酸序列存在于环VIII中。在一些实施方案中,环VIII包含位置571-592,这些位置根据SEQ ID NO:138编号。在一些实施方案中,环VIII包含位置571-599,这些位置根据SEQ ID NO:982编号。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置570、571、572、573、574、575或576之后存在。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列置换位置577(例如,T577)。

[1214] 在一些实施方案中,AAV衣壳变体包含相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于YPAEVVQK(SEQ ID NO:943)的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。

[1215] 在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:2024-2063中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:2024-2063中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,相对于SEQ ID NO:2024-2063中任一个的氨基酸序列包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列的不同氨基酸存在于以下位置中的一处或多处:(i)位置1,其中不同氨基酸为T或L;(ii)位置2,其中不同氨基酸为N、L、K、A、T或P;(iii)位置3,其中不同氨基酸N、K、L、A、Y或S;(iv)位置4,其中不同氨基酸为Q、L、T、S、F、Y、K或A;(v)位置5,其中不同氨基酸为S、H、A、M、Q、T、V或F;(vi)位置6,其中不同氨基酸为S、P、V、A、Q、L、T、N或M;(vii)位置7,其中不同氨基酸为Y、H、S、V、A、L或T;(viii)位置8,其中不同氨基酸为D、P、A、Q、F、L、S、H或M;(ix)位置9,其中不同氨基酸为F、A、L、D或Q;(x)位置10,其中不同氨基酸为T、E、I或S;(xi)位置11,其中不同氨基酸为V、A、N或S;(xii)位置12,其中不同氨基酸为V、L或P;(xiii)位置13,其中不同氨基酸为Q、E或P;(xiv)位置14,其中不同氨基酸为K、N、S或L;(xv)位置15,其中不同氨基酸为T、V、M或L;和/或(xvi)位置16,其中不同氨基酸为A、G或R。在一些实施方案中,所述氨基酸序列存在于环VIII中。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置570、571、572、573、574、575或576之后存在。

[1216] 在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:1632-2023中任一个的氨基酸序列,包含一个、两个或三个但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:1632-2023中任一个的氨基酸序列,包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列。在一些实施方案中,相对于SEQ ID NO:1623-2023中任一个的氨基酸序列包含一个、两个或三个但不超过四个不同氨基酸的氨基酸序列的不同氨基酸存在于以下位置中的一处或多处:(i)位置1,其中不同氨基酸为T、G、N、S、E、L、Y、V或I;(ii)位置2,其中不同氨基酸为D、N、K、E、V、G、R、L、H、F、P、T、A、S、I或Y;(iii)位置3,其中不同氨基酸为Y、N、K、T、W、Q、M、V、C、A、L、F、H、G、R、S或P;(iv)位置4,其中不同氨基酸为H、Q、P、E、R、K、A、S、V、L、T、D、I、G、M或N;(v)位置5,其中不同氨基酸为R、S、K、N、H、G、W、A、P、V、Q、Y、L或F;(vi)位置6,其中不同氨基酸为G、S、F、R、W、H、I、C、M、A、Y、K、N、Q、V、P、E、D、T或L;(vii)位置7,其中不同氨基酸为D、Y、S、I、H、F、P、K、R、G、L、Q、A、M、T、N、V、W、C或E;(viii)位置8,其中不同氨基酸为P、L、Q、T、W、V、G、K、I、Y、N、

H、R、D、S、M、A、F或E;(ix)位置9,其中不同氨基酸为A、R、T、Q、S、M、L、E、K、V、G、D、N、H、F、P或I;(x)位置10,其中不同氨基酸为K、E、Q、H、V、G、R、S、P、I、N、M、A、L、D或T;(xi)位置11,其中不同氨基酸为V、A、E、N、R、L、M、T、Q、S、K、C、G、D、Y、P、H、F或I;(xii)位置12,其中不同氨基酸为V、P、L、S、T、N、A、G、K、R、I、H、E、Q或M;(xiii)位置13,其中不同氨基酸为Q、K、N、A、H、R、T、V、E、I、P、G、S或L;(xiv)位置14,其中不同氨基酸为K、E、I、Y、Q、R、G、D、L、N或S;(xv)位置15,其中不同氨基酸为S、T、N、Q、I、P、E、G、K、M或H;和/或(xvi)位置16,其中不同氨基酸为A、D、L、Y、Q或T。在一些实施方案中,所述氨基酸序列存在于环VIII中。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置570、571、572、573、574、575或576之后存在。

[1217] 在一些实施方案中,AAV衣壳变体包含表1、2A、2B、9和15-20中所提供的任何序列的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:943、1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342、1419、1453、1533、1538、1539、1575、1578、1583-1587、1590、1591-1593、1598-1608或1610-1624中任一个的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:943或2064-2080中任一个的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:2024-2063中任一个的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:1632-2023中任一个的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:943的氨基酸序列。在一些实施方案中,所述氨基酸序列存在于环VIII中。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置570之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置570之后存在,且所述氨基酸序列置换位置571-579(例如,T571、N572、N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置571之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置572-579(例如,N572、N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置571之后存在,且所述氨基酸序列置换位置572-579(例如,N572、N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置572之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置573-579(例如,N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置572之后存在,且所述氨基酸序列置换位置573-579(例如,N573、Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置573之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置574-579(例如,Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置573之后存在,且所述氨基酸序列置换位置574-579(例如,Q574、S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编

号的参考序列,所述氨基酸序列紧接在位置574之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置575-579(例如,S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置574之后存在,且所述氨基酸序列置换位置575-579(例如,S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置575之后存在。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置576-579(例如,S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置575之后存在,且所述氨基酸序列置换位置576-579(例如,S575、S576、T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置576之后存在。在一些实施方案中,所述氨基酸序列置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列置换位置577-579(例如,T577、T578和T579)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置576之后存在,且所述氨基酸序列置换位置577(例如,T577)。在一些实施方案中,相对于根据SEQ ID NO 138编号的参考序列,所述氨基酸序列紧接在位置576之后存在,且所述氨基酸序列置换位置577-579(例如,T577、T578和T579)。

[1218] 在一些实施方案中,AAV衣壳变体包含SEQ ID NO:943或946-966中任一个的氨基酸序列,其中所述氨基酸序列置换相对于SEQ ID NO:138编号的位置577(例如,T577)。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:943或946-966中任一个的氨基酸序列,其中所述氨基酸序列紧接在相对于SEQ ID NO:138编号的位置576之后存在。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:943或946-966中任一个的氨基酸序列,其中所述氨基酸序列紧接在位置576之后存在,且其中所述氨基酸序列置换位置577(例如,T577),这些位置相对于SEQ ID NO:138编号。

[1219] 在一些实施方案中,AAV衣壳变体(例如,本文所述的AAV衣壳变体)包含由以下编码的氨基酸序列:SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含由以下编码的氨基酸序列:SEQ ID NO:944的核苷酸序列,或相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,AAV衣壳变体包含由以下编码的氨基酸序列:相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。

[1220] 在一些实施方案中,编码AAV衣壳变体(例如,本文所述的AAV衣壳变体)的核苷酸序列包含SEQ ID NO:944的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。在一些实施方案中,编码AAV衣壳变体的核酸序列包含相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,编码本文所述的AAV衣壳变体

的核苷酸序列包含相对于SEQ ID NO:944的核苷酸序列,包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个不同核苷酸的核苷酸序列。

[1221] 在一些实施方案中,本文所述的AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列,其中所述氨基酸序列存在于环VIII中。在一些实施方案中,本文所述的AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,所述氨基酸序列紧接在位置576之后存在。在一些实施方案中,本文所述的AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列,其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,YPAEVVQK(SEQ ID NO:943)的氨基酸序列置换位置577(例如,T577)。在一些实施方案中,本文所述的AAV衣壳变体包含YPAEVVQK(SEQ ID NO:943)的氨基酸序列,其中YPAEVVQK(SEQ ID NO:943)的氨基酸序列紧接在位置576之后存在,且其中相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,YPAEVVQK(SEQ ID NO:943)的氨基酸序列置换位置577(例如,T577)。

[1222] 在一些实施方案中,本文所述的AAV衣壳变体包含位置577处的氨基酸Y,且进一步包含PAEVVQK(SEQ ID NO:20)的氨基酸序列,该序列紧接在相对于SEQ ID NO:138编号的位置577之后存在。

[1223] 在一些实施方案中,本文所述的AAV衣壳变体包含位置577处的氨基酸Y和位置578-584处的PAEVVQK(SEQ ID NO:20)的氨基酸序列,这些位置相对于SEQ ID NO:982编号。

[1224] 在一些实施方案中,本文所述的AAV衣壳变体包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)的氨基酸序列,其中所述氨基酸序列存在于环VIII中。在一些实施方案中,AAV衣壳变体包含TNNQSSYPAEVVQKTA(SEQ ID NO:1533)且紧接在相对于SEQ ID NO:138编号的位置570之后存在,其中YPAEVVQK(SEQ ID NO:943)置换相对于SEQ ID NO:138编号的位置577(例如,置换T577)。

[1225] 在一些实施方案中,AAV衣壳变体进一步以下一个、两个、三个或全部:相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置574处非Q的氨基酸(例如,T、S、A、I、L或H),位置575处非S的氨基酸(例如,G、A、L、T或R),和/或位置576处非S的氨基酸(例如,K、L、R、A、Y或T)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的S和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的T、位置575处的S和/或位置576处的L。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的S、位置575处的S和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的S和/或位置576处的R。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的S和/或位置576处的K。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的A、位置575处的G和/或位置576处的A。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的I、位置575处的G和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的A和/或位置576处的S。在一些实施

方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的A、位置575处的S和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的L、位置575处的G和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的S和/或位置576处的T。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的H、位置575处的S和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的L、位置575处的S和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的R和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的S、位置575处的L和/或位置576处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的S、位置575处的L和/或位置576处的Y。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的S、位置575处的A和/或位置576处的T。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置574处的Q、位置575处的T和/或位置576处的S。

[1226] 在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体包含位置574处非Q的氨基酸(例如,S)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体包含位置574处的S。

[1227] 在一些实施方案中,AAV衣壳变体进一步包含以下一个或两个:相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置571处非T的氨基酸(例如,I或N),和/或位置573处非N的氨基酸(例如,T、S或K)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置456处的R。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置571处的T、位置572处的N和/或位置573处的N。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置571处的T、位置572处的N和/或位置573处的T。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置571处的I、位置572处的N和/或位置573处的N。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置571处的T、位置572处的N和/或位置573处的S。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置571处的N、位置572处的N和/或位置573处的N。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置571处的T、位置572处的N和/或位置573处的K。

[1228] 在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置578处非T的氨基酸(例如,P或N)。在一些实施方案中,AAV衣壳变体进一步包含以下一个或两个:相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,位置578处非T的氨基酸(例如,P或N),和/或位置589处非A的氨基酸(例如,D)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步

包含位置578处的T和/或位置588处的A。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置578处的P和/或位置588处的A。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置578处的N和/或位置588处的A。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置578处的T和/或位置579处的D。

[1229] 在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体进一步包含位置577处非T的氨基酸(例如,Y)。在一些实施方案中,相对于根据SEQ ID NO:138的氨基酸序列编号的参考序列,AAV衣壳变体包含位置577处的Y。

[1230] 在一些实施方案中,AAV衣壳变体进一步包含环I、II、IV和/或VI中的修饰(例如,插入、取代和/或缺失)。在一些实施方案中,环I、II、IV、VI和VIII可如以下文献中所述进行鉴定:Govindasamy等人Structurally Mapping the Diverse Phenotype of Adeno-Associated Virus Serotype 4.Journal of Virology.2006年12月80(23):11556-11570;及Govindasamy等人Structural Insights into Adeno-Associated Virus Serotype 5.Journal of Virology.2013年10月87(20):11187-11199;其内容各自由此以引用方式整体并入。

[1231] 在一些实施方案中,可将额外修饰(例如,取代(例如,保守取代)、插入和/或缺失)引入本文所述的AAV衣壳变体中使用野生型AAV5的结构图所确定的位置处,例如以下文献中所述且由此生成的结构图:Govindasamy等人等人Structural Insights into Adeno-Associated Virus Serotype 5.Journal of Virology.2013年10月87(20):11187-11199(其内容由此以引用方式整体并入本文中)或Walters等人“Structure of Adeno-Associated Virus Serotype 5,”Journal of Virology,2004,78(7):3361-3371(其内容由此以引用方式整体并入)。

[1232] 在一些实施方案中,本文所述的AAV衣壳变体包含如以下文献中所述的修饰:Jose等人“High-Resolution Structural Characterization of a New Adenoassociated Virus Serotype 5Antibody Epitope toward Engineering Antibody-Resistant Recombinant Gene Delivery Vectors,”Journal of Virology,2020,93(1):e01394-18;Qian等人“Directed Evolution of AAV Serotype 5for Increased Hepatocyte Transduction and Retained Low Humoral Seroreactivity,”Molecular Therapy: Methods and Clinical Development,2021,20:122-132;Afione等人“Identification and Mutagenesis of the Adeno-Associated Virus 5Sialic Acid Binding Region,”Journal of Virology,2015,89(3):1660-1672;和/或Wang等人“Directed evolution of adeno-associated virus 5capsid enables specific liver tropism,”Mol Ther Nucleic Acids,2022,28:293-306;其各自的内容由此以引用方式整体并入。

[1233] 在一些实施方案中,AAV衣壳变体进一步包括包含SEQ ID NO:138的氨基酸序列的至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过30、20或10个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。在一些实施方案中,AAV衣壳变体进一步包含相对于SEQ ID NO:138的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同氨基酸的氨基酸序列。在一些实施方案中,AAV衣壳变体进一步包含SEQ ID NO:138的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98

或99%)序列同一性的氨基酸序列。

[1234] 在一些实施方案中,AAV衣壳变体进一步包含由SEQ ID NO:137的核苷酸序列编码的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的序列。在一些实施方案中,AAV衣壳变体进一步包含由以下编码的氨基酸序列:相对于SEQ ID NO:137的核苷酸序列,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过30、20或10个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,AAV衣壳变体进一步包含由以下编码的氨基酸序列:相对于SEQ ID NO:137的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同核苷酸的核苷酸序列。

[1235] 在一些实施方案中,编码AAV衣壳变体的核苷酸序列进一步包含SEQ ID NO:137的核苷酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的序列。在一些实施方案中,编码AAV衣壳变体的核苷酸序列进一步包含相对于SEQ ID NO:137的核苷酸序列,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过30、20或10个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,编码AAV衣壳变体的核苷酸序列进一步包含相对于SEQ ID NO:137的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同核苷酸的核苷酸序列。

[1236] 在一些实施方案中,本公开的AAV衣壳变体包含如本文所述的氨基酸序列,例如TTN-002、TTN-003、TTN-004、TTN-005或TTN-006的AAV衣壳变体的氨基酸序列,例如,如表3和表4中所述。在一些实施方案中,本公开的AAV衣壳变体包含如本文所述的氨基酸序列,例如TTN-002的AAV衣壳变体的氨基酸序列,例如,如表3和表4中所述。

[1237] 在一些实施方案中,本文所述的AAV衣壳变体包含VP1、VP2和/或VP3蛋白,其包含本文所述的氨基酸序列,例如TTN-002、TTN-003、TTN-004、TTN-005或TTN-006的AAV衣壳变体的氨基酸序列,例如,如表3和表4中所述。在一些实施方案中,本文所述的AAV衣壳变体包含VP1、VP2和/或VP3蛋白,其包含本文所述的氨基酸序列,例如TTN-002的AAV衣壳变体的氨基酸序列,例如,如表3和表4中所述。

[1238] 在一些实施方案中,本文所述的AAV衣壳变体包含由以下编码的氨基酸序列:如本文所述的核苷酸序列,例如TTN-002的AAV衣壳变体的核苷酸序列,例如,如表3和表5中所述。

[1239] 在一些实施方案中,本公开的编码AAV衣壳变体的多核苷酸或核酸包含本文所述的核苷酸序列,例如TTN-002的AAV衣壳变体的核苷酸序列,例如,如表3和表5中所述。

[1240] 表3.例示性全长衣壳序列

[1241]

名称	VP1 DNA SEQ ID NO:	VP1 氨基酸 SEQ ID NO:	肽 SEQ ID NO:	编码肽的 DNA SEQ ID NO:
TTN-002	984	982	943	944

[1242] 表4.例示性全长衣壳氨基酸序列

[1243]

名称和注释	SEQ ID NO:	氨基酸序列
TTN-002	982	MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQH

[1244]

加下划线的 8 聚体肽，在位置 577 处开始(紧接在位置 576 之后)		QDQARGLVLPGYNYLGPGNGLDRGEPVNRADEVARE HDISYNEQLEAGDNPYLKYNHADAEFQEKLADDTSFG GNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDH FPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPAS SLGADTMSAGGGGPLGDNNQGADGVGNASGDWHCDS TWMGDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSN ANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVF TDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQYGYATL NRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEE VPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQ FNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRA SVSAFATTNRMELEGASYQVPPQPNGMTNNLQGSNTY ALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNR VAYNVGGQMATNNQSS**YPAEVVQK**TAPATGTYNLQE IVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGF GLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQ VTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAP DSTGEYRTTRPIGTRYLTRPL
TTN-002 VP2 (SEQ ID NO: 982 的位置 137-731；加下划线的 8 聚体肽)	738	TAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGS QQLQIPAQPASSLGADTMSAGGGGPLGDNNQGADGVG NASGDWHCDSTWMGDRVVTKSTRTWVLPSYNNHQY REIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPR DWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTI ANNLTSTVQVFTDDDYQLPYVVGNGTEGCLPAFPPQVF TLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTG NNFEFTYNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRF VSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTQG WNLGSGVNRASVSAFATTNRMELEGASYQVPPQPNGM TNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLI TSESETQPVNRVAYNVGGQMATNNQSS**YPAEVVQK**T APATGTYNLQEIVPGSVWMERDVYLQGPIWAKIPETGA HFHPSPAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPV SSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNY NDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL
TTN-002 VP3 (SEQ ID NO: 982 的位置 193-731；加下划线的 8 聚体肽)	739	MSAGGGGPLGDNNQGADGVGNASGDWHCDSTWMGD RVVTKSTRTWVLPSYNNHQYREIKSGSVDGSNANAYF GYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPRS LRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQ LPYVVGNGTEGCLPAFPPQVFTLPQYGYATLNRDNTEN PTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFA PSQNLFKLANPLVDQYLYRFVSTNNTGGVQFNKNLAG RYANTYKNWFPGPMGRTQGWNLGSGVNRASVSAFAT TNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMI FNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVG GQMATNNQSS**YPAEVVQK**TAPATGTYNLQEIVPGSVW MERDVYLQGPIWAKIPETGAHFHPSPAMGGFGLKHPPP MMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEW ELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRT TRPIGTRYLTRPL

[1245]

TTN-003 位置 574-577 处的修饰，及紧接在位置 577 之后存在的 7 聚体插入物(加下划线)	740	MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQH QDQARGLVLPGYNYLGPGNGLDRGEPVNRADEVARE HDISYNEQLEAGDNPYLKYNHADAEFQEKLADDTSFG GNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDH FPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPAS SLGADTMSAGGGGPLGDNNQGADGVGNASGDWHCDS TWMGDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSN ANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVF TDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQYGYATL NRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEE VPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQ FNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRA SVSAFATTNRMELEGASYQVPPQPNGMTNNLQGSNTY ALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNR VAYNVGGQMATNN**AGAYPAEVVQK**TAPATGTYNLQ EIVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGG FGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQ VTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAP DSTGEYRTTRPIGTRYLTRPL
TTN-004 位置 573-575、577 处的修饰，及紧接在位置 577 之后存在的 7 聚体插入物(加下划线)	741	MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQH QDQARGLVLPGYNYLGPGNGLDRGEPVNRADEVARE HDISYNEQLEAGDNPYLKYNHADAEFQEKLADDTSFG GNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDH FPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPAS SLGADTMSAGGGGPLGDNNQGADGVGNASGDWHCDS TWMGDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSN ANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVF TDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQYGYATL NRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEE VPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQ FNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRA SVSAFATTNRMELEGASYQVPPQPNGMTNNLQGSNTY ALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNR VAYNVGGQMATN**STH**S**YPAEVVQK**TAPATGTYNLQE IVPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGF GLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQ VTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAP DSTGEYRTTRPIGTRYLTRPL
TTN-005 位置 574、575、577 处的修饰，及紧接在位置 577 之后存在的 7 聚体插入	742	MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQH QDQARGLVLPGYNYLGPGNGLDRGEPVNRADEVARE HDISYNEQLEAGDNPYLKYNHADAEFQEKLADDTSFG GNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDH FPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPAS SLGADTMSAGGGGPLGDNNQGADGVGNASGDWHCDS TWMGDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSN ANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVF TDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQYGYATL

[1246]

物(加下划线)		NRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEE VPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQ FNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRA SVSAFATTNRMELEGASYQVPPQPNGMTNNLQGSNTY ALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNR VAYNVGGQMATNN**IGSYPAEVVQK**TAPATGTYNLQEI VPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGF GLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQ VTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAP DSTGEYRTTRPIGTRYLTRPL
TTN-006 位置 577 处的修饰，及紧接在位置 577 之后存在的 7 聚体插入物(加下划线)	743	MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQH QDQARGLVLPGYNYLGPGNGLDRGEPVNRADEVARE HDISYNEQLEAGDNPYLKYNHADAEFQEKLADDTSFG GNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDH FPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPAS SLGADTMSAGGGGPLGDNNQGADGVGNASGDWHCDS TWMGDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSN ANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVF TDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQYGYATL NRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEE VPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQ FNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNRA SVSAFATTNRMELEGASYQVPPQPNGMTNNLQGSNTY ALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNR VAYNVGGQMATNNQSS**YPAERSHK**TAPATGTYNLQEI VPGSVWMERDVYLQGPIWAKIPETGAHFHPSPAMGGF GLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQ VTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAP DSTGEYRTTRPIGTRYLTRPL

[1247] 表5.例示性全长衣壳核酸序列

[1248]

名称和注释	SEQ ID NO:	NT 序列
TTN-002 编码 8 聚体肽的核苷酸序列(加下划线)	984	ATGTCTTTTGTTGATCACCCTCCAGATTGGTTGGA AGAAGTTGGTGAAGGTCTTCGCGAGTTTTTGGGCC TTGAAGCGGGCCCACCGAAACCAAAACCCAATCA GCAGCATCAAGATCAAGCCCGTGGTCTTGTGCTGC CTGGTTATAACTATCTCGGACCCGGAAACGGTCTC GATCGAGGAGAGCCTGTCAACAGGGCAGACGAGG TCGCGCGAGAGCACGACATCTCGTACAACGAGCA GCTTGAGGCGGGAGACAACCCCTACCTCAAGTAC AACCACGCGGACGCCGAGTTTCAGGAGAAGCTCG CCGACGACACATCCTTCGGGGGAAACCTCGGAAA GGCAGTCTTTCAGGCCAAGAAAAGGGTTCTCGAA CCTTTTGGCCTGGTTGAAGAGGGTGCTAAGACGGC CCCTACCGGAAAGCGGATAGACGACCACTTTCCA AAAAGAAAGAAGGCCCGGACCGAAGAGGACTCC

[1249]

		AAGCCTTCCACCTCGTCAGACGCCGAAGCTGGAC CCAGCGGATCCCAGCAGCTGCAAATCCCAGCCCA ACCAGCCTCAAGTTTGGGAGCTGATACAATGTCTG CGGGAGGTGGCGGCCCATTGGGCGACAATAACCA AGGTGCCGATGGAGTGGGCAATGCCTCGGGAGAT TGGCATTGCGATTCCACGTGGATGGGGGACAGAG TCGTCACCAAGTCCACCCGAACCTGGGTGCTGCCC AGCTACAACAACCACCAGTACCGAGAGATCAAAA GCGGCTCCGTCGACGGAAGCAACGCCAACGCCTA CTTTGGATACAGCACCCCCTGGGGGTACTTTGACT TTAACCGCTTCCACAGCCACTGGAGCCCCCGAGAC TGGCAAAGACTCATCAACAACTACTGGGGCTTCA GACCCCGGTCCCTCAGAGTCAAAATCTTCAACATT CAAGTCAAAGAGGTCACGGTGCAGGACTCCACCA CCACCATCGCCAACAACCTCACCTCCACCGTCCAA GTGTTTACGGACGACGACTACCAGCTGCCCTACGT CGTCGGCAACGGGACCGAGGGATGCCTGCCGGCC TTCCCTCCGCAGGTCTTTACGCTGCCGCAGTACGG TTACGCGACGCTGAACCGCGACAACACAGAAAAT CCCACCGAGAGGAGCAGCTTCTTCTGCCTAGAGTA CTTTCCCAGCAAGATGCTGAGAACGGGCAACAAC TTTGAGTTTACCTACAACTTTGAGGAGGTGCCCTT CCACTCCAGCTTCGCTCCCAGTCAGAACCTCTTCA AGCTGGCCAACCCGCTGGTGGACCAGTACTTGTAC CGCTTCGTGAGCACAAATAACACTGGCGGAGTCC AGTTCAACAAGAACCTGGCCGGGAGATACGCCAA CACCTACAAAAACTGGTTCCCGGGGCCCATGGGC CGAACCCAGGGCTGGAACCTGGGCTCCGGGGTCA ACCGCGCCAGTGTCAGCGCCTTCGCCACGACCAAT AGGATGGAGCTCGAGGGCGCGAGTTACCAGGTGC CCCCGCAGCCGAACGGCATGACCAACAACCTCCA GGGCAGCAACACCTATGCCCTGGAGAACACTATG ATCTTCAACAGCCAGCCGGCGAACCCGGGCACCA CCGCCACGTACCTCGAGGGCAACATGCTCATCACC AGCGAGAGCGAGACGCAGCCGGTGAACCGCGTGG CGTACAACGTCGGCGGGCAGATGGCCACCAACAA CCAGAGCTCT**TATCCGGCGGAGGTGGTGCAGAA** **G**actgcccccgcgaccggcaCGTACAACCTCCAGGAAATCG TGCCCGGCAGCGTGTGGATGGAGAGGGACGTGTA CCTCCAAGGACCCATCTGGGCCAAGATCCCAGAG ACGGGGGCGCACTTTCACCCCTCTCCGGCtATGGG CGGATTCGGACTCAAACACCCACCGCCCATGATG CTCATCAAGAACACGCCTGTGCCCGGAAATATCA CCAGCTTCTCGGACGTGCCCGTCAGCAGCTTCATC ACCCAGTACAGCACCGGGCAGGTCACCGTGGAGA TGGAGTGGGAGCTCAAGAAGGAAAACTCCAAGAG GTGGAACCCAGAGATCCAGTACACAAACAACTAC AACGACCCCCAGTTTGTGGACTTTGCCCCGGACAG CACCGGGGAATACAGAACCACCAGACCTATCGGA ACCCGATACCTTACCCGACCCCTTTAA

[1250] 在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的氨基酸序

列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少90%序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少95%序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少96%序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少97%序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少98%序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其具有至少99%序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含相对于SEQ ID NO:982的氨基酸序列,包含至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过30、20或10个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:982的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同氨基酸的氨基酸序列。

[1251] 在一些实施方案中,在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:740-743中任一个的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含相对于SEQ ID NO:740-743中任一个的氨基酸序列,包含至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过30、20或10个修饰(例如,取代(例如,保守取代)、插入或缺失)的氨基酸序列。在一些实施方案中,AAV衣壳变体包含相对于SEQ ID NO:740-743中任一个的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同氨基酸的氨基酸序列。

[1252] 在一些实施方案中,本文所述的AAV衣壳变体包含由以下编码的氨基酸序列:SEQ ID NO:984的核苷酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的核苷酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含由以下编码的氨基酸序列:相对于SEQ ID NO:984的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同核苷酸的核苷酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含由以下编码的氨基酸序列:相对于SEQ ID NO:984的核苷酸序列,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过30、20或10个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1253] 在一些实施方案中,编码本文所述的AAV衣壳变体的核苷酸序列包含SEQ ID NO:984的核苷酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的核苷酸序列。在一些实施方案中,编码本文所述的AAV衣壳变体的核苷酸序列包含SEQ ID NO:984的核苷酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的核苷酸序列。在一些实施方案中,编码本文所述的AAV衣壳变体的核苷酸序列包含相对于SEQ ID NO:984的核苷酸序列,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过30、20或10个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,编码本文所述的AAV衣壳变体的核苷酸序列包含相对于SEQ ID NO:984的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同核苷酸的核苷酸序列。

在一些实施方案中,编码本文所述的AAV衣壳变体的核酸序列为密码子优化的。

[1254] 在一些实施方案中,本文所述的AAV衣壳变体包含VP1、VP2、VP3蛋白或其组合。在一些实施方案中,AAV衣壳变体包含对应于SEQ ID NO:982的位置137-731(例如,VP2)的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的序列。在一些实施方案中,所述AAV衣壳蛋白包含对应于SEQ ID NO:982的位置193-731(例如,VP3)的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的序列。在一些实施方案中,AAV衣壳变体包含对应于SEQ ID NO:982的位置1-731(例如,VP1)的氨基酸序列,或与其具有至少70%(例如,至少约80、85、90、95、96、97、98或99%)序列同一性的氨基酸序列。

[1255] 在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:739的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:738的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:982的氨基酸序列,或与其至少95%(例如,至少96、97、98或99%)同一的氨基酸序列。

[1256] 在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:739的氨基酸序列(例如,VP3)。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:738的氨基酸序列(例如,VP2)。在一些实施方案中,AAV衣壳变体包含SEQ ID NO:982的氨基酸序列(例如,VP1)。

[1257] 在一些实施方案中,相对于包含SEQ ID NO:138的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性。

[1258] 在一些实施方案中,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性。

[1259] 在一些实施方案中,相对于包含SEQ ID NO:139的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对CNS细胞或组织(例如,脑细胞、脑组织、脊髓细胞或脊髓组织)具有增加的趋向性。

[1260] 在一些实施方案中,本文所述的AAV衣壳变体转导脑区,例如颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团和/或小脑。在一些实施方案中,如与SEQ ID NO:139的参考序列相比,转导水平为至少.5、2.2、2.4、2.5、2.6、2.7、3.0、3.2、3.5、3.7、4.0、4.2、4.5、4.7、4.9、5、10、15、20、25、30或35倍大。

[1261] 在一些实施方案中,与SEQ ID NO:138的参考序列相比,本文所述的AAV衣壳变体在脑中富集至少约2、3、4、5、6、7、8、9、10、12、15、17、20、25、30、35、40、45、50、55、60或65倍。在一些实施方案中,与SEQ ID NO:138的参考序列相比,本文所述的AAV衣壳变体在脑中富集至少约10、12、15、17、20、25、30、35、40、45、50、55、60、61、62、63、64或65倍。

[1262] 在一些实施方案中,与SEQ ID NO:138的参考序列相比,本文所述的AAV衣壳变体在至少两个至三个物种(例如,非人灵长类动物和啮齿动物(例如,大鼠和/或小鼠)物种)的脑中富集。在一些实施方案中,与SEQ ID NO:138或982的参考序列相比,本文所述的AAV衣壳变体在至少两个至三个物种(例如,非人灵长类动物和啮齿动物(例如,大鼠和/或小鼠)物种)的脑中富集至少约2、3、4、5、6、7、8、9、10、12、15、17、20、25、30、35、40、45、50、60、70、

80、90或100倍。在一些实施方案中,至少两个至三个物种为食蟹猕猴、西非绿猴、普通狨、大鼠和/或小鼠(例如,BALB/c小鼠)。

[1263] 在一些实施方案中,与SEQ ID NO:982的参考序列相比,本文所述的AAV衣壳变体在脑中富集约5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、100、125、150、175、200或225倍。

[1264] 在一些实施方案中,本文所述的AAV衣壳变体将提高水平的病毒基因组递送至脑区。在一些实施方案中,如与SEQ ID NO:139的参考序列相比,病毒基因组的水平提高至少1.5、2.2、2.4、2.5、2.6、2.7、3.0、3.2、3.5、3.7、4.0、4.2、4.5、4.7、4.9或5倍。在一些实施方案中,所述脑区包含中脑区(例如,海马体或丘脑)和/或脑干。

[1265] 在一些实施方案中,本文所述的AAV衣壳变体将提高水平的有效载荷递送至脑区。在一些实施方案中,如与SEQ ID NO:139的参考序列相比,有效载荷的水平提高至少20、25、30、35倍。在一些实施方案中,所述脑区包含颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑或其组合。

[1266] 在一些实施方案中,与SEQ ID NO:139的参考序列相比,本文所述的AAV衣壳变体在脊髓区中富集至少约3、3.5、4.0、4.5、5、5.0、6.0或6.5倍。在一些实施方案中,所述脊髓区包含脊髓颈区、脊髓腰区、脊髓胸区或其组合。

[1267] 在一些实施方案中,相对于背根神经节(DRG)中的转导,本文所述的AAV衣壳变体显示脑区中的优先转导。在一些实施方案中,相对于肝中的转导,本文所述的AAV衣壳变体显示脑区中的优先转导。

[1268] 在一些实施方案中,本文所述的AAV衣壳变体能够转导神经元细胞。

[1269] 在一些实施方案中,本文所述的AAV衣壳变体能够转导非神经元细胞,例如神经胶质细胞(例如,少突胶质细胞或星形胶质细胞)。在一些实施方案中,AAV衣壳变体能够转导神经元细胞和非神经元细胞,例如神经胶质细胞(例如,少突胶质细胞或星形胶质细胞)。在一些实施方案中,非神经元细胞为神经胶质细胞(例如,少突胶质细胞或星形胶质细胞)。

[1270] 在一些实施方案中,相对于包含SEQ ID NO:138的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对心脏细胞或心脏组织(例如,心房或心室的心脏细胞或心脏组织)具有增加的趋向性。在一些实施方案中,相对于包含SEQ ID NO:139的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对心脏细胞或心脏组织(例如,心房或心室的心脏细胞或心脏组织)具有增加的趋向性。

[1271] 在一些实施方案中,本文所述的AAV衣壳变体将提高水平的有效载荷递送至心脏区域。在一些实施方案中,如与SEQ ID NO:139的参考序列相比,有效载荷的水平提高至少1.5、2或2.5倍。

[1272] 在一些实施方案中,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对心脏细胞或心脏组织(例如,心房或心室的心脏细胞或心脏组织)具有增加的趋向性。在一些实施方案中,与SEQ ID NO:982的参考序列相比,AAV衣壳变体在心脏中富集约4、5、10、15、20、25、30、35、40、45或50倍。

[1273] 在一些实施方案中,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对肌肉细胞或组织具有增加的趋向性。在一些实施方案中,肌肉区域包含股四头肌和/或横膈膜肌肉区域。在一些实施方案中,与SEQ ID NO:982的参考

序列相比,AAV衣壳变体在肌肉中富集至少约2、3、4、5、10、15、20、25、30或35倍。

[1274] 在一些实施方案中,相对于包含SEQ ID NO:982的氨基酸序列的参考序列的趋向性,本文所述的AAV衣壳变体对肝细胞或肝组织具有增加的趋向性。在一些实施方案中,与SEQ ID NO:982的参考序列相比,AAV衣壳变体在肝中富集至少约10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、115、120、130、140、150、160、170、180、185或190倍。

[1275] 在一些实施方案中,本文所述的AAV衣壳变体对肝具有降低的趋向性。在一些实施方案中,AAV衣壳变体包含修饰(例如,取代、插入或缺失),其导致肝中降低的趋向性(例如,脱靶)和/或活性。在一些实施方案中,将肝中降低的趋向性与不包含所述修饰的其他方面相似的衣壳(例如,野生型衣壳多肽)进行比较。在一些实施方案中,所述的AAV衣壳变体包含修饰(例如,取代、插入或缺失),所述修饰导致以下特性中的一个或多个:(1)肝中降低的趋向性;(2)肝中的脱靶表达;(3)肝中降低的活性;和/或(4)降低的与半乳糖的结合。在一些实施方案中,将特性(1)-(3)中的任一个或全部的降低与不包含所述修饰的其它方面相似的AAV衣壳变体进行比较。在一些实施方案中,AAV衣壳变体(例如,具有肝中降低的趋向性的AAV衣壳变体)包含以下一个或多个:位置581处非A、G、K、M、N、Q、R、S和/或T的氨基酸;位置582处非A、C、H、I、K、S、T和/或V的氨基酸;位置583处非A、G、H、K、M、N、Q、R、S、T和/或V的氨基酸;位置584处非L、M、P、Q、R、T和/或W的氨基酸;位置585处非F、H、I、K、M、T和/或Y的氨基酸;位置586处非E、G、H、L、M、N、Q、T和/或W的氨基酸;位置587处非A、C、G、H、L、M、R和/或S的氨基酸;位置588处非A、C、D、F、G、H、M、Q、S、V、W和/或Y的氨基酸;和/或位置589处非A、C、E、G、H、M、N、P、Q、S、V和/或W的氨基酸,这些位置均相对于SEQ ID NO:138编号。

[1276] 在一些实施方案中,本公开的AAV衣壳变体为分离的,例如重组的。在一些实施方案中,编码本公开的AAV衣壳多肽(例如,AAV衣壳变体)的多核苷酸为分离的,例如重组的。

[1277] 本文还提供编码任何上述AAV衣壳变体的多核苷酸序列,及包含所述序列的AAV颗粒、载体和细胞。

[1278] AAV血清型和衣壳

[1279] 在一些实施方案中,本公开的AAV颗粒可包含衣壳蛋白或其变体(任何天然或重组AAV血清型)。AAV血清型的特征可能不同,诸如但不限于包装、趋向性、转导和免疫原性型态。尽管不希望受理论束缚,但据信在一些实施方案中,AAV衣壳蛋白(例如,AAV衣壳变体)可调节特定组织中的AAV颗粒趋向性。

[1280] 在一些实施方案中,本文所述的AAV衣壳变体允许静脉内施用之后的血脑障壁穿透。在一些实施方案中,AAV衣壳变体允许静脉内施用、聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合之后的血脑障壁穿透。在一些实施方案中,AAV衣壳变体允许向脑区的分布增加。在一些实施方案中,脑区包含颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、小脑、额叶皮层、感觉皮层、运动皮层、齿状核、小脑皮层、大脑皮层、脑干或其组合。在一些实施方案中,相对于背根神经节(DRG)中的转导,AAV衣壳变体允许脑区中的优先转导。在一些实施方案中,相对于肝中的转导,AAV衣壳变体允许脑区中的优先转导。在一些实施方案中,AAV衣壳变体允许神经元细胞中的转导。在一些实施方案中,AAV衣壳变体允许非神经元细胞,例如神经胶质细胞(例如,星形胶质细胞、少突胶质细胞或其组合)中的转导。在一些

实施方案中,AAV衣壳变体允许神经元细胞和非神经元细胞,例如神经胶质细胞(例如,星形胶质细胞、少突胶质细胞或其组合)中的转导。

[1281] 在一些实施方案中,AAV衣壳变体允许向脊髓区的分布增加。在一些实施方案中,脊柱区包含脊髓颈区、脊髓胸区和/或脊髓腰区。

[1282] 在一些实施方案中,AAV衣壳变体允许向心脏区域的分布增加。

[1283] 在一些实施方案中,AAV衣壳变体适合肌肉内施用和/或肌肉纤维转导。在一些实施方案中,AAV衣壳变体允许向肌肉区域的分布增加。在一些实施方案中,肌肉区域包含心脏肌肉、股四头肌、横膈膜肌肉区域或其组合。

[1284] 在一些实施方案中,AAV衣壳变体允许向肝区域的分布增加。

[1285] 在一些实施方案中,本文所述的AAV衣壳包含如以下文献中所述的修饰:Jose等人High-Resolution Structural Characterization of aNew Adenoassociated Virus Serotype 5Antibody Epitope toward Engineering Antibody-Resistant Recombinant Gene Delivery Vectors.Journal of Virology.2019年1月93(1):e01394-18;Qian等人Directed Evolution of AAV Serotype 5for Increased Hepatocyte Transduction and Retained Low Humoral Seroreactivity.Molecular Therapy:Methods&Clinical Development.2020年10月20:122-132;Afione等人Identification and Mutagenesis of the Adeno-Associated Virus 5Sialic Binding Region.Journal of Virology.2015年2月89(3):1660-1672;其内容各自由此以引用方式整体并入。

[1286] 在一些实施方案中,用于本文所述的AAV VP1衣壳蛋白(例如,衣壳变体)的转译的起始密码子可为如美国专利号US8163543中所述的CTG、TTG或GTG,该专利的内容以引用方式整体并入本文中。

[1287] 本公开涉及由衣壳(Cap)基因编码的结构衣壳蛋白(包括VP1、VP2和VP3)。这些衣壳蛋白形成诸如AAV的病毒载体的外部蛋白质结构壳(例如,衣壳)。由Cap多核苷酸合成的VP衣壳蛋白一般包括甲硫氨酸作为肽序列中的第一个氨基酸(Met1),其与相应Cap核苷酸序列中的起始密码子(AUG或ATG)相关。然而,常见的是,第一个甲硫氨酸(Met1)残基或一般任何第一个氨基酸(AA1)在多肽合成之后或期间由蛋白质加工酶(诸如Met-氨基肽酶)裂解出来。此"Met/AA剪切"过程通常与多肽序列中的第二个氨基酸(例如,丙氨酸、缬氨酸、丝氨酸、苏氨酸等)的相应乙酰化相关。Met剪切通常发生于VP1和VP3衣壳蛋白中,但也可发生于VP2衣壳蛋白中。

[1288] 在Met/AA剪切不完整的情况下,可产生包含病毒衣壳的一种或多种(一种、两种或三种)VP衣壳蛋白的混合物,其中一些可包括Met1/AA1氨基酸(Met+/AA+)且其中一些可由于Met/AA剪切而缺乏Met1/AA1氨基酸(Met-/AA-)。关于衣壳蛋白中的Met/AA剪切的进一步论述,参见Jin等人Direct Liquid Chromatography/Mass Spectrometry Analysis for Complete Characterization of Recombinant Adeno-Associated Virus Capsid Proteins.Hum Gene Ther Methods.2017年10月28(5):255-267;Hwang等人N-Terminal Acetylation of Cellular Proteins Creates Specific Degradation Signals.Science.2010年2月19日327(5968):973-977;其内容各自以引用方式整体并入本文中。

[1289] 根据本公开,对衣壳蛋白(例如,AAV衣壳变体)的提及不限于剪切的(Met-/AA-)或

未剪切的(Met+/AA+),且在上下文中可指独立衣壳蛋白、包含衣壳蛋白和/或编码、描述、产生或导致本公开的衣壳蛋白的多核苷酸序列(或其片段)的混合物的病毒衣壳。对衣壳蛋白或衣壳多肽(诸如VP1、VP2或VP2)的直接提及还可包含包括Met1/AA1氨基酸(Met+/AA+)的VP衣壳蛋白以及由于Met/AA剪切而缺乏Met1/AA1氨基酸(Met-/AA-)的相应VP衣壳蛋白。

[1290] 此外,根据本公开,对分别包含或编码一种或多种包括Met1/AA1氨基酸(Met+/AA+)的衣壳蛋白的特定SEQ ID NO:(蛋白质抑或核酸)的提及应理解为教示缺乏Met1/AA1氨基酸之VP衣壳蛋白,因为在对序列进行审查时,显而易见任何仅缺乏第一个所列氨基酸(无论为否为Met1/AA1)的序列。

[1291] 作为非限制性实例,对736个氨基酸的长度且包括由AUG/ATG起始密码子编码的“Met1”氨基酸(Met+)的VP1多肽序列的提及也可理解为教示735个氨基酸的长度且不包括所述736个氨基酸的Met+序列的“Met1”氨基酸(Met-)的VP1多肽序列。作为第二非限制性实例,对736个氨基酸的长度且包括由任何NNN起始密码子编码的“AA1”氨基酸(AA1+)的VP1多肽序列的提及还可理解为教示735个氨基酸的长度且不包括所述736个氨基酸的AA1+序列的“AA1”氨基酸(AA1-)的VP1多肽序列。

[1292] 对由VP衣壳蛋白形成的病毒衣壳的提及(诸如对特定AAV衣壳血清型的提及)可并入包括Met1/AA1氨基酸(Met+/AA1+)的VP衣壳蛋白、由于Met/AA1剪切而缺乏Met1/AA1氨基酸(Met-/AA1-)的相应VP衣壳蛋白及其组合(Met+/AA1+和Met-/AA1-)。

[1293] 作为非限制性实例,AAV衣壳血清型可包括VP1(Met+/AA1+)、VP1(Met-/AA1-)或VP1(Met+/AA1+)与VP1(Met-/AA1-)的组合。AAV衣壳血清型还可包括VP3(Met+/AA1+)、VP3(Met-/AA1-)或VP3(Met+/AA1+)与VP3(Met-/AA1-)的组合;且还可包括VP2(Met+/AA1)和VP2(Met-/AA1-)的相似的任选地选用的组合。

[1294] 额外AAV序列

[1295] 在一些实施方案中,AAV衣壳变体包含表1、2A、2B、9或15-20中所提供的任何氨基酸序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15或16个连续氨基酸,所述氨基酸紧接在位置570、571、572、573、574、575或576之后,这些位置相对于SEQ ID NO:138编号。

[1296] 在一些实施方案中,AAV衣壳变体包含表1、2A、2B、9或15-22中所提供的任何氨基酸序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15或16个连续氨基酸,所述氨基酸紧接在位置570、571、572、573、574、575、576或577之后,这些位置相对于SEQ ID NO:138编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。在一些实施方案中,表1、2A、2B、9或15-22中所提供的任何氨基酸序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15或16个连续氨基酸置换位置T571、N572、N573、Q574、S575、S576、T577、T578和/或A579中的至少一个、两个、三个、四个、五个、六个、七个、八个或全部,这些位置根据SEQ ID NO:138编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。在一些实施方案中,表1、2A、2B、9或15-22中所提供的任何氨基酸序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15或16个连续氨基酸置换位置T577,该位置根据SEQ ID NO:138

编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。在一些实施方案中,AAV衣壳变体包含位置T571、N572、N573、Q574、S575、S576、T577、T578和/或A579中的一个、两个、三个、四个、五个、六者、七者、八者或全部位置处非野生型(例如,原生)氨基酸的氨基酸,这些位置根据SEQ ID NO:138编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。在一些实施方案中,AAV衣壳变体包含位置T577处非野生型(例如,原生)氨基酸的氨基酸,该位置根据SEQ ID NO:138编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。在一些实施方案中,AAV衣壳变体包含位置T571、N572、N573、Q574、S575、S576、T577、T578和/或A579中的一个、两个、三个、四个、五个、六者、七者、八者或全部位置处的修饰(例如,取代),这些位置根据SEQ ID NO:138编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。在一些实施方案中,AAV衣壳变体包含位置T577处的修饰(例如,取代),该位置根据SEQ ID NO:138编号或对应于任何其他AAV血清型(例如,AAV1、AAV2、AAV3、AAV3b、AAV4、AAV6、AAV7、AAV8、AAV9、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、SEQ ID NO:139、PHP.N、PHP.B或如WO 2021/230987(其内容由此以引用方式整体并入)表6中所提供的AAV血清型)中的等效位置。

[1297] 在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少90%同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少95%同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少96%同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少97%同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少98%同一性的氨基酸序列。在一些实施方案中,本文所述的AAV衣壳变体包含SEQ ID NO:138的氨基酸序列,或与其具有至少99%同一性的氨基酸序列。在一些实施方案中,AAV衣壳多肽或AAV衣壳变体包含相对于SEQ ID NO:138的氨基酸序列,包含至少一个、两个或三个修饰(例如,取代(例如,保守取代))但不超过30、20或10个修饰(例如,取代(例如,保守取代))的氨基酸序列。在一些实施方案中,AAV衣壳多肽或AAV衣壳变体包含相对于SEQ ID NO:138的氨基酸序列,包含至少一个、两个或三个但不超过30、20或10个不同氨基酸的氨基酸序列。在一些实施方案中,AAV衣壳多肽或AAV衣壳变体包含由以下编码的氨基酸序列:SEQ ID NO:137的

核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。在一些实施方案中,编码AAV衣壳多肽或AAV衣壳变体的核苷酸序列包含SEQ ID NO:137的核苷酸序列,或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的核苷酸序列。

[1298] 在一些实施方案中,本文所述的AAV衣壳变体包含所述氨基酸序列或由以下任一个的核苷酸序列编码:US7427396的SEQ ID NO:1或2;US20030138772的SEQ ID NO:114;US20150315612的SEQ ID NO:199;或US20160289275A1的SEQ ID NO:13、14、16、17、19、20、22、23、25、26、28、29、31、32、34、35、37、38、40、41、43或44(各自的内容以引用方式整体并入本文中),或与其大体上同一(例如,具有至少70%、75%、80%、85%、90%、92%、95%、97%、98%或99%序列同一性)的序列。

[1299] 在一些实施方案中,本文所述的AAV衣壳变体包含位置569(例如,M569V)、652(例如,D652A)、362(例如,T362M)、359(例如,Q359D)、350(例如,E350Q)、533(例如,P533S)、585(例如,Y585V)、587(例如,L587T)、581(例如,A581T)、582(例如,T582A)、584(例如,T584A)或其组合处的修饰(例如,取代),全部位置均相对于SEQ ID NO:138编号。

[1300] 在一些实施方案中,本文所述的AAV衣壳变体包含位置581至589中的一个或多个处来自野生型AAV5序列(例如,SEQ ID NO:138的氨基酸序列)的氨基酸,这些位置相对于SEQ ID NO:138编号。在一些实施方案中,AAV衣壳变体包含以下1、2、3、4、5、6、7、8个或全部:位置581处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置581处的氨基酸A);位置582处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置582处的氨基酸T);位置583处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置583处的氨基酸G);位置584处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置584处的氨基酸T);位置585处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置585处的氨基酸Y);位置586处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置586处的氨基酸N);位置587处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置587处的氨基酸L);位置588处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置588处的氨基酸Q);和/或位置589处来自野生型AAV5序列(例如,SEQ ID NO:138)的氨基酸(例如,包含位置589处的氨基酸E)。

[1301] 在某些实施方案中,本文所述的AAV衣壳不包含位置581处的T、位置582处的A、位置584处的A、位置585处的V、位置585处的T、位置569处的V、位置652处的A、位置362处的M、位置359处的Q、位置350处的Q、位置533处的S或其组合,全部位置均相对于SEQ ID NO:138编号。

[1302] 在一些实施方案中,本文所述的AAV衣壳不包含位置581-589(根据SEQ ID NO:138编号)处的修饰(例如,取代),其中所述修饰具有WO 2021/242909的表2、7、8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、39、41、43、45、47、49、51、53、55、57、59、61、63、65、67、69或71-86中所提供的任何序列的氨基酸序列。

[1303] 在任何本文所述的实施方案中,可如下鉴定相对于SEQ ID NO:138编号的位置:提供参考序列与查询序列的比对,其中参考序列为SEQ ID NO:138,且鉴定查询序列中对应于

参考序列中的位置的位置所对应的残基。

[1304] 表6.AAV序列

[1305]

血清型	SEQ ID NO:	序列
AAV5 WT (氨基酸)	138	MSFVDHPPDWLEEVGEGLREFLGLEAGPPKPKPNQQ HQDQARGLVLPGYNYLGPGNGLDRGEPVNRADEV AREHDISYNEQLEAGDNPYLKYNHADAEFQEKLAD DTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPT GKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQ QLQIPAQPASSLGADTMSAGGGGPLGDNNQGADGV GNASGDWHCDSTWMGDRVVTKSTRTWVLPSYNNH QYREIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSH WSPRDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQ DSTTTIANNLTSTVQVFTDDDYQLPYVVGNGTEGCL PAFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEY FPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFKLA NPLVDQYLYRFVSTNNTGGVQFNKNLAGRYANTYK NWFPGPMGRTQGWNLGSGVNRASVSAFATTNRME LEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQ PANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQ MATNNQSSTTAPATGTYNLQEIVPGSVWMERDVYL QGPIWAKIPETGAHFHPSPAMGGFGLKHPPPMMLIK NTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELK KENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTT RPIGTRYLTRPL
AAV5 WT VP2(氨基酸; SEQ ID NO: 138 的位置 137-724)	744	TAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPS GSQQLQIPAQPASSLGADTMSAGGGGPLGDNNQGA DGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSY NNHQYREIKSGSVDGSNANAYFGYSTPWGYFDFNR FHSHWSPRDWQRLINNYWGFRPRSLRVKIFNIQVKE VTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVGNGT EGCLPAFPPQVFTLPQYGYATLNRDNTENPTERSSFF CLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNL FKLANPLVDQYLYRFVSTNNTGGVQFNKNLAGRYA NTYKNWFPGPMGRTQGWNLGSGVNRASVSAFATT NRMELEGASYQVPPQPNGMTNNLQGSNTYALENTM IFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYN VGGQMATNNQSSTTAPATGTYNLQEIVPGSVWMER DVYLQGPIWAKIPETGAHFHPSPAMGGFGLKHPPPM MLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEME WELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTG EYRTTRPIGTRYLTRPL
AAV5	750	MSAGGGGPLGDNNQGADGVGNASGDWHCDSTWM

[1306]

WT VP3 (氨基酸；SEQ ID NO: 138 的位置 193-724)		GDRVVTKSTRTWVLPSYNNHQYREIKSGSVDGSNA NAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYW GFRPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQ VFTDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQYGY ATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFT YNFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTN NTGGVQFNKNLAGRYANTYKNWFPGPMGRTQGW NLGSGVNRASVSAFATTNRMELEGASYQVPPQPNG MTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGN MLITSESETQPVNRVAYNVGGQMATNNQSSTTAPAT GTYNLQEIVPGSVWMERDVYLQGPIWAKIPETGAHF HPSPAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPV SSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTN NYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL
AAV5 WT (DNA)	137	atgtcttttgttgatcaccctccagattggttggaagaagttggtgaaggtcttcgcga gtttttgggccttgaagcgggcccaccgaaaccaaaaccccaatcagcagcatcaa gatcaagcccgtggtcttgtgctgcctggttataactatctcggacccggaaacggt ctcgatcgaggagagcctgtcaacagggcagacgaggtcgcgcgagagcacga catctcgtacaacgagcagcttgaggcgggagacaaccctacctcaagtacaac cacgcggacgccgagtttcaggagaagctcgccgacgacacatccttcgggggga aacctcggaaaggcagtctttcaggccaagaaaagggttctcgaacctttggcct ggttgaagagggtgctaagacggcccctaccggaaagcggatagacgaccactt tccaaaaagaaagaaggctcggaccgaagaggactccaagccttccacctcgtc agacgccgaagctggacccagcggatcccagcagctgcaaatcccagcccaac cagcctcaagtttgggagctgatacaatgtctgcgggaggtggcggcccattggg cgacaataaccaaggtgccgatggagtgggcaatgcctcgggagattggcattgc gattccacgtggatgggggacagagtcgtcaccaagtccacccgaacctgggtg ctgcccagctacaacaaccaccagtaccgagagatcaaaagcggctccgtcgac ggaagcaacgccaacgcctactttggatacagcacccctgggggtactttgactt taaccgcttccacagccactggagcccccgagactggcaaagactcatcaacaac tactggggcttcagaccccggtccctcagagtcaaaatcttcaacattcaagtcaaa gaggtcacggtgcaggactccaccaccaccatcgccaacaacctcacctccacc gtccaagtgtttacggacgacgactaccagctgccctacgtcgtcggcaacggga ccgagggatgcctgccggccttccctccgcaggtctttacgctgccgcagtacggt tacgcgacgctgaaccgcgacaacacagaaaatcccaccgagaggagcagcttc ttctgcctagagtactttcccagcaagatgctgagaacgggcaacaactttgagttta cctacaactttgaggaggtgcccttccactccagcttcgctcccagtcagaacctgtt caagctggccaacccgctggtggaccagtacttgtaccgcttcgtgagcacaaata acactggcggagtccagttcaacaagaacctggccgggagatacgccaacacct acaaaaactggttcccggggcccatgggccgaacccagggctggaacctgggct ccggggtcaaccgcgccagtgtcagcgccttcgccacgaccaataggatggagc tcgagggcgcgagttaccaggtgcccccgcagccgaacggcatgaccaacaac ctccagggcagcaacacctatgccctggagaacactatgatcttcaacagccagc cggcgaacccgggcaccaccgccacgtacctcgagggcaacatgctcatcacca gcgagagcgagacgcagccggtgaaccgcgtggcgtacaacgtcggcgggca gatggccaccaacaaccagagctcTACTACTGCCCCCGCGACC GGCACGTACAACCTCCAGGAAATCGTGCCCGGCA GCGTGTGGATGGAGAGGGACGTGTACCTCCAAGG ACCCATCTGGGCCAAGATCCCAGAGACGGGGGCG CACTTTCACCCCTCTCCGGCCATGGGCGGATTCGG

[1307]

		ACTCAAACACCCACCGCCCATGATGCTCATCAAG AACACGCCTGTGCCCGGAAATATCACCAGCTTCTC GGACGTGCCCGTCAGCAGCTTCATCACCCAGTACA GCACCGGGCAGGTCACCGTGGAGATGGAGTGGGA GCTCAAGAAGGAAAACTCCAAGAGGTGGAACCCA GAGATCCAGTACACAAACAACTACAACGACCCCC AGTTTGTGGACTTTGCCCCGGACAGCACCGGGGA ATACAGAACCACCAGACCTATCGGAACCCGATAC CTTACCCGACCCCTTTAA
AAV9/hu.14 WT (氨基酸)	139	MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKA NQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNAA DAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQE RLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKT APGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFG QTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAP VADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTR TWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPW GYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKL FNIQVKEVTDNNGVKTIANNLTSTVQVFTDSDYQLP YVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAV GRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSY AHSQSLDRLMNPLIDQYLYYLSKTINGSGQNQQTLK FSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNN NSEFAWPGASSWALNGRNSLMNPGPAMASHKEGE DRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKT TNPVATESYGQVATNHQSAQAQAQTGWVQNQGILP GMVWQDRDVYLQGPIWAKIPHTDGNFHPSPLMGGF GMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYS TGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSNNV EFAVNTEGVYSEPRPIGTRYLTRNL

[1308] AAV颗粒的病毒基因组

[1309] 在一些实施方案中,包含本文所述的AAV衣壳变体的如本文所述的AAV颗粒可用于将病毒基因组递送至组织(例如,CNS、心脏、肝和/或肌肉)。在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒可用于将病毒基因组递送至组织或细胞,例如CNS、DRG、心脏、肝或肌肉细胞或组织。在一些实施方案中,本公开的AAV颗粒为重组AAV颗粒。在一些实施方案中,本公开的AAV颗粒为分离的AAV颗粒。

[1310] 病毒基因组可编码任何有效载荷,诸如但不限于多肽(例如,治疗性多肽)、抗体、酶、RNAi剂和/或基因编辑系统的组分。在一个实施方案中,本文所述的AAV颗粒用于在静脉内递送之后将有效载荷递送至CNS的细胞。在另一实施方案中,本文所述的AAV颗粒用于在静脉内递送之后将有效载荷递送至DRG的细胞。在一些实施方案中,本文所述的AAV颗粒用于在静脉内递送之后将有效载荷递送至心脏的细胞。在一些实施方案中,本文所述的AAV颗粒用于在静脉内递送之后将有效载荷递送至肌肉(例如,心肌)的细胞。在一些实施方案中,本文所述的AAV颗粒用于在静脉内递送之后将有效载荷递送至肝的细胞。

[1311] 在一些实施方案中,包含如本文所述的AAV衣壳变体的AAV颗粒的病毒基因组包括包含编码有效载荷的转基因的核苷酸序列。在一些实施方案中,病毒基因组包含反向末端

重复序列(ITR)。在一些实施方案中,病毒基因组包含两个ITR序列,一个序列位于病毒基因组的5'端(例如,相对于所编码的有效载荷的5')且一个序列位于病毒基因组的3'端(例如,相对于所编码的有效载荷的3')。在一些实施方案中,AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)的病毒基因组可包含调节组件(例如,启动子)、非转译区(UTR)、miR结合位点、多腺苷酸化序列(polyA)、填充(filler/stuffer)序列、内含子和/或接头序列,例如,用于增强转基因表达。

[1312] 在一些实施方案中,所述病毒基因组组分被选择和/或被工程改造以在靶组织(例如,CNS(例如,脑、脊髓或两者)、心脏、肝和/或肌肉组织)中表达有效载荷。

[1313] 病毒基因组组分:反向末端重复序列(ITR)

[1314] 在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒包括包含ITR和编码有效载荷的转基因的病毒基因组。在一些实施方案中,病毒基因组包含两个ITR。在一些实施方案中,这两个ITR侧接5'和3'端编码有效载荷的核苷酸序列。在一些实施方案中,这些ITR充当复制起点,其包含复制的识别位点。在一些实施方案中,这些ITR包含可互补且对称排列的序列区域。在一些实施方案中,并入如本文所述的病毒基因组中的ITR可包含天然存在的多核苷酸序列或重组衍生的多核苷酸序列。

[1315] 在一些实施方案中,ITR可来自与选自任何已知血清型或其变体的衣壳多肽(例如,衣壳变体)相同的血清型。在一些实施方案中,ITR可具有与所述衣壳不同的血清型。在一些实施方案中,病毒基因组包含两个ITR序列区域,其中ITR具有彼此相同的血清型。在一些实施方案中,病毒基因组包含两个ITR序列区域,其中ITR具有不同血清型。非限制性实例包括具有与所述衣壳相同的血清型的ITR中的零个、一个或两个。在一些实施方案中,AAV颗粒的病毒基因组的两个ITR皆为AAV2 ITR。

[1316] 独立地,每个ITR可为约100至约150个核苷酸的长度。ITR可为约100-105个核苷酸的长度、106-110个核苷酸的长度、111-115个核苷酸的长度、116-120个核苷酸的长度、121-125个核苷酸的长度、126-130个核苷酸的长度、131-135个核苷酸的长度、136-140个核苷酸的长度、141-145个核苷酸的长度或146-150个核苷酸的长度。在一些实施方案中,ITR为140-142个核苷酸的长度。ITR长度的非限制性实例为102、105、130、140、141、142、145个核苷酸的长度。

[1317] 病毒基因组组分:启动子

[1318] 在一些实施方案中,本文所述的AAV颗粒的病毒基因组包含至少一种增强有效载荷靶标特异性和表达的组件(参见,例如Powell等人Viral Expression Cassette Elements to Enhance Transgene Target Specificity and Expression in Gene Therapy,2015;其内容以引用方式整体并入本文中)。增强有效载荷靶标特异性和表达的组件的非限制性实例包括启动子、内源性miRNA、转录后调节组件(PRE)、多腺苷酸化(PolyA)信号序列及上游增强子(USE)、CMV增强子和内含子。

[1319] 在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒包括包含核酸的病毒基因组,所述核酸包含编码有效载荷的转基因,其中转基因可操作地连接至启动子。在一些实施方案中,启动子为物种特异性启动子、诱导型启动子、组织特异性启动子或细胞周期特异性启动子(例如,如Parr等人,Nat.Med.3:1145-9(1997)中所述的启动子;其内容以引用方式整体并入本文中)。

[1320] 在一些实施方案中,启动子可为天然存在或非天然存在的。启动子的非限制性实例包括源自病毒、植物、哺乳动物或人的那些。在一些实施方案中,启动子可为源自人细胞或系统的那些。在一些实施方案中,启动子可为截短或突变的,例如启动子变体。

[1321] 在一些实施方案中,启动子为普遍存在的启动子,例如能够在多种组织中表达。在一些实施方案中,启动子为人类延伸因子1α-亚单位(EF1α)启动子、巨细胞病毒(CMV)即刻早期增强子和/或启动子、鸡β-肌动蛋白(CBA)启动子及其衍生物CAG、β葡糖醛酸酶(GUSB)启动子或泛素C(UBC)启动子。在一些实施方案中,启动子为细胞或组织特异性启动子,例如能够在中枢或外周神经系统的组织或细胞、其中细胞内的靶向区域(例如,额叶皮层)和/或子集(例如,兴奋性神经元)中表达。在一些实施方案中,启动子为能够在兴奋性神经元(例如,谷氨酸激导性)、抑制性神经元(例如,GABA激导性)、交感或副交感神经系统的神经元、感觉神经元、背根神经节的神经元、运动神经元或神经系统的支持细胞(诸如小神经胶质细胞、神经胶质细胞、星形胶质细胞、少突胶质细胞和/或许旺细胞)中表达有效载荷的细胞类型特异性启动子。

[1322] 在一些实施方案中,启动子为肝特异性启动子(例如,hAAT、TBG)、骨骼肌特异性启动子(例如,结蛋白、MCK、C512)、B细胞启动子、单核细胞启动子、白血球启动子、巨噬细胞启动子、胰脏腺泡细胞启动子、内皮细胞启动子、肺组织启动子和/或心脏或心血管启动子(例如,αMHC、cTnT和CMV-MLC2k)。

[1323] 在一些实施方案中,启动子为用于中枢神经系统的组织或细胞中的有效载荷表达的组织特异性启动子。在一些实施方案中,启动子为突触蛋白(Syn)启动子、谷氨酸囊泡转运蛋白(VGLUT)启动子、囊泡GABA转运蛋白(VGAT)启动子、小清蛋白(PV)启动子、钠通道Na_v1.8启动子、酪氨酸羟化酶(TH)启动子、胆碱乙酰转移酶(ChaT)启动子、甲基-CpG结合蛋白2(MeCP2)启动子、Ca^{2+}/钙调蛋白依赖性蛋白激酶II(CaMKII)启动子、代谢型谷氨酸受体2(mGluR2)启动子、神经丝轻链(NFL)或重链(NFH)启动子、神经元特异性烯醇化酶(NSE)启动子、β-球蛋白微基因nβ2启动子、前脑啡肽原(PPE)启动子、脑啡肽(Enk)启动子和兴奋性氨基酸转运蛋白2(EAAT2)启动子或其片段。在一些实施方案中,启动子为能够在星形胶质细胞中表达的细胞类型特异性启动子,例如神经胶质纤维酸性蛋白(GFAP)启动子和EAAT2启动子或其片段。在一些实施方案中,启动子为能够在少突胶质细胞中表达的细胞类型特异性启动子,例如髓磷脂碱性蛋白(MBP)启动子或其片段。

[1324] 在一些实施方案中,启动子为GFAP启动子。在一些实施方案中,启动子为突触蛋白(syn/syn1)启动子或其片段。

[1325] 在一些实施方案中,启动子包含胰岛素启动子或其片段。

[1326] 在一些实施方案中,本文所述的病毒基因组的启动子(例如,包含于包含本文所述的AAV衣壳变体的AAV颗粒内)包含EF-1α启动子或其变体,例如,如表40中所提供。在一些实施方案中,EF-1α启动子包含SEQ ID NO:2100、2101、2103、2104、2108、2109、2111-2120中任一个的核苷酸序列,或以下任一个:表8中所提供的核苷酸序列,相对于SEQ ID NO:2100、2101、2103、2104、2108、2109、2111-2120的核苷酸序列包含至少一个、两个、三个、四个、五个、六个或七个但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列,或以下任一个:表8中所提供的核苷酸序列,或与SEQ ID NO:2100、2101、2103、2104、2108、2109、2111-2120中任一个具有至少70%(例如,80、85%、90%、95%、96%、97%、98%或99%)序列同一性的

核苷酸序列,或表8中所提供的任一核苷酸序列。

[1327] 表8.例示性启动子变体

[1328]

描述	序列	SEQ ID NO:
EF1a 启动子(内含子加下划线)	CGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCAC ATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTC GGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGG TAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTT TTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAG TAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCC GCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCG GGCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCT TGAATTACTTCCACCTGGCTGCAGTACGTGATTCTTGA TCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCG AGGCCTTGCGCTTAAGGAGCCCCTTCGCCTCGTGCTTG AGTTGAGGCCTGGCCTGGGCGCTGGGGCCGCCGCGTG CGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTT CGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTG CTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAAAT GCGGGCCAAGATCTGCACACTGGTATTTCGGTTTTTGG GGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGC ACATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCAC CGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCC TGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCCC CGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGT TGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCT GCAGGGAGCTCAAAATGGAGGACGCGGCGCTCGGGAG AGCGGGCGGGTGAGTCACCCACACAAAGGAAAAGGG CCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACG GAGTACCGGGCGCCGTCCAGGCACCTCGATTAGTTCTC GAGCTTTTGGAGTACGTCGTCTTTAGGTTGGGGGGAGG GGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTG GAGACTGAAGTTAGGCCAGCTTGGCACTTGATGTAATT CTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTT CATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTC TTCCATTTCAGGTGTCGTGA	2100
miniEF1a	GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCC CCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGG TGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGT GATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGG GGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAAC GTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACGCGT AAG	2101
启动子变体 1	GCATG	

[1329]

启动子变体 2	GGTGGAGAAGAGCATG	2103
启动子变体 3	GTCATCACTGAGGTGGAGAAGAGCATG	2104
启动子变体 4	CGTGAG	
启动子变体 5	GT	
启动子变体 6	GCTCCGGT	
启动子变体 19	GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAG	2108
启动子变体 20	GCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACGC	2109
启动子变体 7	GTAAG	
启动子变体 8	GTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACGCGTAAG	2111
启动子变体 9	GCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACGCGTAAG	2112
启动子变体 10	CGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACGCGTAAG	2113
启动子变体 11	CGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGG	2114

[1330]

	TAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTT TTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAG TAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCC GCCAGAACACAG	
启动子变体 12	GCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAG CGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAG GGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCG CGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTC CGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAA GTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGG TTTGCCGCCAGAACACGCGTAAG	2115
启动子变体 13	GCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAG CGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAG GGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCG CGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTC CGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAA GTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGG TTTGCCGCCAGAACACAG	2116
启动子变体 14	GGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCCCGTC AGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAA GTTGGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAG AGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCG TGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAA CCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTT TCGCAACGGGTTTGCCGCCAGAACACGCGTAAG	2117
启动子变体 15	GGTGGAGAAGAGCATGCGTGAGGCTCCGGTGCCCGTC AGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAA GTTGGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAG AGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCG TGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAA CCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTT TCGCAACGGGTTTGCCGCCAGAACACAG	2118
启动子变体 16	GTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTC CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACA GTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAA CCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGA AAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGG GTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGT GAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAC GCGTAAG	2119
启动子变体 18	GTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTC CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACA GTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAA CCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGA AAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGG GTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGT GAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACAC AG	2120

[1331] 病毒基因组组分:非转译区(UTR)

[1332] 在一些实施方案中,基因的野生型非转译区(UTR)被转录,但未被转译。一般来说,

5'UTR在转录起始位点处起始且在起始密码子处终止,且3'UTR紧接在终止密码子之后开始且继续直至转录终止信号。

[1333] 典型地在特定靶器官(例如,CNS组织、肌肉或DRG)的大量表达基因中发现的特征可被工程改造至UTR中以提高稳定性和蛋白质产生。作为非限制性实例,通常在脑中表达的mRNA(例如,亨廷顿蛋白)的5'UTR可用于本文所述的AAV颗粒的病毒基因组中以增强中枢神经系统的神经元细胞或其他细胞中的表达。

[1334] 不希望受理论束缚,野生型5′非转译区(UTR)包括在转译起始中发挥作用的特征。通常已知,Kozak序列参与核糖体起始多种基因的转译的过程,通常包括于5'UTR中。Kozak序列具有共有CCR(A/G)CCAUGG,其中R为起始密码子(ATG)上游三个碱基的嘌呤(腺嘌呤或鸟嘌呤),其后为另一'G'。

[1335] 在一个实施方案中,病毒基因组中的5'UTR包括Kozak序列。

[1336] 在一个实施方案中,病毒基因组中的5'UTR不包括Kozak序列。

[1337] 虽然不希望受理论束缚,但已知野生型3′UTR具有包埋于其中的腺苷和尿苷延伸段。这些富AU印记在具有高周转率的基因中尤其普遍。基于其序列特征和功能特性,这些富AU组件(ARE)可分为三类(Chen等人,1995,其内容以引用方式整体并入本文中):I类ARE,诸如但不限于c-Myc和MyoD,在富U区域内含有数个分散的AUUUA模体拷贝。II类ARE(诸如但不限于GM-CSF和TNF-a)具有两个或两个以上重叠的UUAUUUA(U/A)(U/A)九聚体。III类ARE(诸如但不限于c-Jun和肌细胞生成素)未明确定义。这些富U区域不含AUUUA模体。已知大多数与ARE结合的蛋白质会使信使不稳定,而ELAV家族的成员(尤其为HuR)已被记录提高mRNA的稳定性。HuR与所有三个类别的ARE结合。将HuR特异性结合位点工程改造至核酸分子的3′UTR中将导致HuR结合且因此导致信使在体内稳定。

[1338] 3′UTR富AU组件(ARE)的引入、移除或修饰可用于调节多核苷酸的稳定性。当工程改造特定多核苷酸(例如,病毒基因组的有效载荷区域)时,可引入ARE的一个或多个拷贝以使多核苷酸不太稳定且由此减少转译且减少所得蛋白质的产生。同样,可鉴定ARE且进行移除或突变以提高细胞内稳定性且因此增加所得蛋白质的转译和产生。

[1339] 在一个实施方案中,病毒基因组的3'UTR可包括用于模板化添加poly-A尾的oligo(dT)序列。

[1340] 在一个实施方案中,病毒基因组可包括至少一个miRNA种子、结合位点或完整序列。微小RNA(或miRNA或miR)为19-25个核苷酸的非编码RNA,其与核酸靶标的位点结合且通过降低核酸分子稳定性或通过抑制转译来下调基因表达。在一些实施方案中,微小RNA序列包含种子区域,例如成熟微小RNA的位置2-8区域中的序列,其具有与所述核酸的miRNA靶序列完全或部分互补的Watson-Crick序列。

[1341] 在一个实施方案中,病毒基因组可被工程改造以包括、改变或移除至少一个miRNA结合位点、完整序列或种子区域。

[1342] 来自本领域中已知的任何基因的任何UTR均可并入所述AAV颗粒的病毒基因组中。这些UTR或其部分可以与其中选出UTR或其部分的基因中相同的取向置放,或其可改变取向或位置。在一个实施方案中,所述AAV颗粒的病毒基因组中所用的UTR可被倒置、缩短、延长,与本领域中已知的一个或多个其他5′UTR或3′UTR一起制成。如本文所用,与UTR相关的术语"改变(altered)"意谓UTR相对于参考序列已以某种方式发生改变。例如,相对于野生型或

原生UTR,3′或5′UTR可如上文所教示通过取向或位置的改变来改变,或可通过包括额外核苷酸、核苷酸缺失、核苷酸交换或转座来改变。

[1343] 在一个实施方案中,所述AAV颗粒的病毒基因组包含至少一个并非野生型UTR变体的人工UTR。

[1344] 在一个实施方案中,所述AAV颗粒的病毒基因组包含选自转录本家族的UTR,这些转录本的蛋白质共享共同功能、结构、特征或特性。

[1345] 病毒基因组组分:多腺苷酸化序列

[1346] 本文所述的AAV颗粒的病毒基因组(例如,包含本文所述的AAV衣壳变体的AAV颗粒)可包含多腺苷酸化序列。在一些实施方案中,所述AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)的病毒基因组包含在编码有效载荷的核苷酸序列的3'端与3'ITR的5'端之间的多腺苷酸化序列。

[1347] 病毒基因组组分:内含子

[1348] 在一些实施方案中,如本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的病毒基因组包含增强有效载荷靶标特异性和表达组件(参见,例如Powell等人Viral Expression Cassette Elements to Enhance Transgene Target Specificity and Expression in Gene Therapy,Discov.Med,2015,19(102):49-57;其内容以引用方式整体并入本文中),诸如内含子。内含子的非限制性实例包括MVM(67-97bp)、F.IX截短的内含子1(300bp)、β-球蛋白SD/免疫球蛋白重链剪接受体(250bp)、腺病毒剪接供体/免疫球蛋白剪接受体(500bp)、SV40晚期剪接供体/剪接受体(19S/16S)(180bp)以及杂合腺病毒剪接供体/IgG剪接受体(230bp)。

[1349] 病毒基因组组分:填充序列

[1350] 在一些实施方案中,本文所述的AAV颗粒的病毒基因组包含改善包装效率和表达的组件,诸如填充(stuffer/filler)序列。填充序列的非限制性实例包括白蛋白和/或α-1抗胰蛋白酶。可操纵任何已知的病毒、哺乳动物或植物序列以用作填充序列。

[1351] 在一些实施方案中,填充序列可为约100-3500个核苷酸的长度。填充序列可具有约100、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900或3000个核苷酸的长度。

[1352] 病毒基因组组分:miRNA

[1353] 在一些实施方案中,病毒基因组包含编码miRNA的序列,以减少有效载荷在组织或细胞(例如,DRG(背根神经节)或其他神经节的神经元,诸如交感或副交感神经系统的那些)中的表达。在一些实施方案中,可在病毒基因组中编码miRNA(例如,miR183、miR182和/或miR96)以调节(例如,减少)病毒基因组在DRG神经元中的表达。作为另一非限制性实例,可在病毒基因组中编码miR-122miRNA以调节(例如,减少)病毒基因组在肝中的表达。在一些实施方案中,可在病毒基因组中编码miRNA(例如,miR-142-3p)以调节(例如,减少)病毒基因组在造血谱系的细胞或组织中的表达,该造血谱系包括例如免疫细胞(例如,抗原呈递细胞或APC,包括树突细胞(DC)、巨噬细胞和B淋巴细胞)。在一些实施方案中,可在病毒基因组中编码miRNA(例如,miR-1)以调节(例如,减少)病毒基因组在心脏的细胞或组织中的表达。

[1354] 病毒基因组组分:miR结合位点

[1355] 可通过引入组织或细胞特异性调节序列,例如启动子、增强子、微小RNA结合位点(例如,脱靶位点)来增强本文所公开的AAV病毒颗粒的组织或细胞特异性表达。不希望受理论束缚,据信所编码的miR结合位点可基于相应内源性微小RNA(miRNA)或相应受控外源性miRNA在组织或细胞(例如,非靶向细胞或组织)中的表达来调节(例如,预防、阻遏或以其他方式抑制)相关基因在本文所公开的病毒基因组上的表达。在一些实施方案中,miR结合位点调节(例如,减少)由本文所述的AAV颗粒的病毒基因组编码的有效载荷在其中表达相应mRNA的细胞或组织中的表达。

[1356] 在一些实施方案中,本文所述的AAV颗粒的病毒基因组包含编码微小RNA结合位点(例如,脱靶位点)的核苷酸序列。在一些实施方案中,本文所述的AAV颗粒的病毒基因组包含编码miR结合位点、微小RNA结合位点系列(miR BSs)或其反向补体的核苷酸序列。

[1357] 在一些实施方案中,编码miR结合位点系列或miR结合位点的核苷酸序列位于病毒基因组的3'-UTR区(例如,相对于编码有效载荷的核苷酸序列的3')中,例如在polyA序列、病毒基因组的5'-UTR区(例如,相对于编码有效载荷的核苷酸序列的5')或两者之前。

[1358] 在一些实施方案中,所编码的miR结合位点系列包含miR结合位点(miR BS)的至少1-5个拷贝,例如至少1-3、2-4、3-5、1、2、3、4、5个或更多个拷贝。在一些实施方案中,所有拷贝均一致,例如包含相同miR结合位点。在一些实施方案中,所编码的miR结合位点系列内的miR结合位点为连续的且未由间隔子分开。在一些实施方案中,所编码的miR结合位点系列内的miR结合位点由间隔子(例如,非编码序列)分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子编码序列或其反向补体包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,具有至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的核苷酸序列。

[1359] 在一些实施方案中,所编码的miR结合位点系列包含miR结合位点(miR BS)的至少1-5个拷贝,例如至少1-3、2-4、3-5、1、2、3、4、5个或更多个拷贝。在一些实施方案中,这些拷贝中的至少1、2、3、4、5者或全部为不同的,例如包含不同miR结合位点。在一些实施方案中,所编码的miR结合位点系列内的miR结合位点为连续的且未由间隔子分开。在一些实施方案中,所编码的miR结合位点系列内的miR结合位点由间隔子(例如,非编码序列)分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸)的长度。在一些实施方案中,间隔子包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,具有至少一个、两个或三个修饰(例如,取代(例如,保守取代)、插入)但不超过四个修饰(例如,取代(例如,保守取代)、插入或缺失)的核苷酸序列。

[1360] 在一些实施方案中,所编码的miR结合位点与宿主细胞中的miR大体上同一(例如,至少70%、75%、80%、85%、90%、95%、99%或100%同一)。在一些实施方案中,所编码的miR结合位点包含与宿主细胞中的miR的至少1、2、3、4或5个错配或不超过6、7、8、9或10个错配。在一些实施方案中,错配的核苷酸为邻接的。在一些实施方案中,错配的核苷酸为非邻

接的。在一些实施方案中,错配的核苷酸出现在miR结合位点的种子区域结合序列外部,诸如在miR结合位点的一端或两端。在一些实施方案中,miR结合位点与宿主细胞中的miR 100%一致。

[1361] 在一些实施方案中,编码miR结合位点的核苷酸序列与宿主细胞中的miR大体上互补(例如,至少70%、75%、80%、85%、90%、95%、99%或100%互补)。在一些实施方案中,编码miR结合位点的核苷酸序列的互补序列包含与宿主细胞中的miR的至少1、2、3、4或5个错配或不超过6、7、8、9或10个错配。在一些实施方案中,错配的核苷酸为邻接的。在一些实施方案中,错配的核苷酸为非邻接的。在一些实施方案中,错配的核苷酸出现在miR结合位点的种子区域结合序列外部,诸如在miR结合位点的一端或两端。在一些实施方案中,所编码的miR结合位点与宿主细胞中的miR 100%互补。

[1362] 在一些实施方案中,所编码的miR结合位点或序列区域为至少约10至约125个核苷酸的长度,例如至少约10至50个核苷酸、10至100个核苷酸、50至100个核苷酸、50至125个核苷酸或100至125个核苷酸的长度。在一些实施方案中,所编码的miR结合位点或序列区域为至少约7至约28个核苷酸的长度,例如至少约8-28个核苷酸、7-28个核苷酸、8-18个核苷酸、12-28个核苷酸、20 -26个核苷酸、22个核苷酸、24个核苷酸或26个核苷酸的长度,且任选地包含至少一个与miRNA(例如,miR122、miR142、miR183或miR1)的种子序列互补(例如,完全或部分互补)的连续区域(例如,7或8个核苷酸)。

[1363] 在一些实施方案中,所编码的miR结合位点与在肝或肝细胞中表达的miR(诸如miR122)互补(例如,完全或部分互补)。在一些实施方案中,所编码的miR结合位点或所编码的miR结合位点系列包含miR122结合位点序列。在一些实施方案中,所编码的miR122结合位点包含ACAAACACCATTGTCACACTCCA(SEQ ID NO:4673)的核苷酸序列,或相对于SEQ ID NO:4673的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,插入、缺失或取代)的核苷酸序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,病毒基因组包含所编码的miR122结合位点(例如,所编码的miR122结合位点系列)的至少2、3、4或5个拷贝,任选地其中所编码的miR122结合位点系列包含以下核苷酸序列:ACAAACACCATTGTCACACTCCACACAAACACCATTGTCACAC TCCACACAAACACCATTGTCACACTCCA(SEQ ID NO:4674),或相对于SEQ ID NO:4674的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,至少两个所编码的miR122结合位点直接连接,例如无间隔子。在其他实施方案中,至少两个所编码的miR122结合位点由间隔子(例如,1、2、3、4、5、6、7、8、9或10个核苷酸的长度)分开,间隔子位于两个或两个以上连续的所编码的miR122结合位点序列之间。在实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个)的长度。在一些实施方案中,间隔子编码序列或其反向补体包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,所编码的miR结合位点系列包含miR122结合位点的至少3-5个拷贝(例如,4个拷贝),具有或不具

有间隔子,其中间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的长度。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1364] 在一些实施方案中,所编码的miR结合位点与在心脏中表达的miR互补(例如,完全或部分互补)。在实施方案中,所编码的miR结合位点或所编码的miR结合位点系列包含miR-1结合位点。在一些实施方案中,所编码的miR-1结合位点包含ATACATACTTCTTTACATTCCA(SEQ ID NO:4679)的核苷酸序列,相对于SEQ ID NO:4679的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或具有至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代(例如,保守取代)、插入或缺失)但不超过十个修饰(例如,取代(例如,保守取代)、插入或缺失)的核苷酸序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,病毒基因组包含所编码的miR-1结合位点(例如,所编码的miR-1结合位点系列)的至少2、3、4或5个拷贝。在一些实施方案中,所编码的miR-1结合位点的至少2、3、4或5个拷贝(例如,2或3个拷贝)为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1365] 在一些实施方案中,所编码的miR结合位点与在造血谱系中表达的miR互补(例如,完全或部分互补),该造血谱系包括免疫细胞(例如,抗原呈递细胞或APC,包括树突细胞(DC)、巨噬细胞和B淋巴细胞)。在一些实施方案中,与在造血谱系中表达的miR互补的所编码的miR结合位点包含公开于例如US2018/0066279中的核苷酸序列,该专利的内容以引用方式整体并入本文中。

[1366] 在实施方案中,所编码的miR结合位点或所编码的miR结合位点系列包含miR-142-3p结合位点序列。在一些实施方案中,所编码的miR-142-3p结合位点包含TCCATAAAGTAGGAAACACTACA(SEQ ID NO:4675)的核苷酸序列,相对于SEQ ID NO:4675的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的核苷酸序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,病毒基因组包含所编码的miR-142-3p结合位点(例如,所编码的miR-142-3p结合位点系列)的至少2、3、4或5个拷贝。在一些实施方案中,所编码的miR-142-3p结合位点的至少2、3、4或5个拷贝(例如,2或3个拷贝)为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例

如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1367] 在一些实施方案中,所编码的miR结合位点与在DRG(背根神经节)神经元中表达的miR(例如,miR183、miR182和/或miR96结合位点)互补(例如,完全互补或部分互补)。在一些实施方案中,所编码的miR结合位点与在DRG神经元中表达的miR互补,包含公开于例如WO2020/132455中的核苷酸序列,其内容以引用方式整体并入本文中。

[1368] 在一些实施方案中,所编码的miR结合位点或所编码的miR结合位点系列包含miR183结合位点序列。在一些实施方案中,所编码的miR183结合位点包含AGTGAATTCTACCAGTGCCATA(SEQ ID NO:4676)的核苷酸序列,或相对于SEQ ID NO:4676的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代(例如,保守取代)、插入或缺失)的核苷酸序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,与种子序列互补的序列对应于所编码的miR-183结合位点序列的双下划线。在一些实施方案中,病毒基因组包含所编码的miR183结合位点(例如,所编码的miR183结合位点)的至少2、3、4或5个拷贝(例如,至少2或3个拷贝)。在一些实施方案中,所编码的miR183结合位点的至少2、3、4或5个拷贝(例如,2或3个拷贝)为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。

[1369] 在一些实施方案中,所编码的miR结合位点或所编码的miR结合位点系列包含miR182结合位点序列。在一些实施方案中,所编码的miR182结合位点包含AGTGTGAGTTCTACCATTGCCAAA(SEQ ID NO:4677)的核苷酸序列,相对于SEQ ID NO:4677的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,病毒基因组包含所编码的miR182结合位点(例如,所编码的miR182结合位点系列)的至少2、3、4或5个拷贝。在一些实施方案中,所编码的miR182结合位点的至少2、3、4或5个拷贝(例如,2或3个拷贝)为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。

[1370] 在某些实施方案中,所编码的miR结合位点或所编码的miR结合位点系列包含miR96结合位点序列。在一些实施方案中,所编码的miR96结合位点包含

AGCAAAAATGTGCTAGTGCCAAA(SEQ ID NO:4678)的核苷酸序列,相对于SEQ ID NO:4678的核苷酸序列,具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、至少95%、至少99%或100%序列同一性,或包含至少一个、两个、三个、四个、五个、六个或七个修饰(例如,取代、插入或缺失)但不超过十个修饰(例如,取代、插入或缺失)的序列,例如,其中修饰可导致所编码的miR结合位点与相应miRNA之间的错配。在一些实施方案中,病毒基因组包含所编码的miR96结合位点(例如,所编码的miR96结合位点系列)的至少2、3、4或5个拷贝。在一些实施方案中,所编码的miR96结合位点的至少2、3、4或5个拷贝(例如,2或3个拷贝)为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。

[1371] 在一些实施方案中,所编码的miR结合位点系列包含miR122结合位点、miR-1、miR142结合位点、miR183结合位点、miR182结合位点、miR 96结合位点或其组合。在一些实施方案中,所编码的miR结合位点系列包含miR122结合位点、miR142结合位点、miR183结合位点、miR182结合位点、miR 96结合位点或其组合的至少2、3、4或5个拷贝。在一些实施方案中,至少两个所编码的miR结合位点经直接连接,例如无间隔子。在其他实施方案中,至少两个所编码的miR结合位点由间隔子(例如,1、2、3、4、5、6、7、8、9或10个核苷酸的长度)分开,间隔子位于两个或两个以上连续的所编码的miR结合位点序列之间。在实施方案中,间隔子为至少约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的长度。在一些实施方案中,间隔子编码序列或其反向补体包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1372] 在一些实施方案中,所编码的miR结合位点系列包含miR-1、miR122结合位点、miR142结合位点、miR183结合位点、miR182结合位点、miR96结合位点中的至少两个、三个、四个、五个或全部的至少2-5个拷贝(例如,2或3个拷贝),其中该系列内的每个miR结合位点均为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含至少一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1373] 在一些实施方案中,所编码的miR结合位点系列包含miR-122结合位点与miR-1结合位点的组合的至少2-5个拷贝(例如,2或3个拷贝),其中该系列内的每个miR结合位点均为连续的(例如,未由间隔子分开)或由间隔子分开。在一些实施方案中,间隔子为约1至6个核苷酸或约5至10个核苷酸(例如,约7-8个核苷酸或约8个核苷酸)的核苷酸长度。在一些实施方案中,间隔子序列包含以下一个或多个:(i)GGAT;(ii)CACGTG;(iii)GCATGC,或(i)-

(iii)中的一个或多个的重复。在一些实施方案中,间隔子包含GATAGTTA的核苷酸序列,或相对于GATAGTTA的核苷酸序列,包含一个、两个或三个修饰(例如,取代、插入或缺失)但不超过四个修饰(例如,取代、插入或缺失)的核苷酸序列。

[1374] 基因组大小

[1375] 在一个实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可包含单链或双链病毒基因组。病毒基因组的大小可为小、中等、大或最大大小。如上文所述,病毒基因组可包含启动子和polyA尾。

[1376] 在一个实施方案中,病毒基因组可为小单链病毒基因组。小单链病毒基因组可为2.1至3.5kb的大小,诸如但不限于约2.1、2.2、2.3、2.4、2.5、2.6、2.7、2.8、2.9、3.0、3.1、3.2、3.3、3.4和3.5kb的大小。

[1377] 在一个实施方案中,病毒基因组可为小双链病毒基因组。小双链病毒基因组可为1.3至1.7kb的大小,诸如但不限于约1.3、1.4、1.5、1.6和1.7kb的大小。

[1378] 在一个实施方案中,病毒基因组可为中等单链病毒基因组。中等单链病毒基因组可为3.6至4.3kb的大小,诸如但不限于约3.6、3.7、3.8、3.9、4.0、4.1、4.2和4.3kb的大小。

[1379] 在一个实施方案中,病毒基因组可为中等双链病毒基因组。中等双链病毒基因组可为1.8至2.1kb的大小,诸如但不限于约1.8、1.9、2.0和2.1kb的大小。

[1380] 在一个实施方案中,病毒基因组可为大单链病毒基因组。大单链病毒基因组可为4.4至6.0kb之大小,例诸如但不限于约4.4、4.5、4.6、4.7、4.8、4.9、5.0、5.1、5.2、5.3、5.4、5.5、5.6、5.7、5.8、5.9和6.0kb的大小。

[1381] 在一个实施方案中,病毒基因组可为大双链病毒基因组。大双链病毒基因组可为2.2至3.0kb的大小,诸如但不限于约2.2、2.3、2.4、2.5、2.6、2.7、2.8、2.9和3.0kb的大小。

[1382] 有效载荷

[1383] 在一些实施方案中,本公开的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)包括包含编码有效载荷的核酸的病毒基因组。在一些实施方案中,所编码的有效载荷为RNAi剂或多肽。本公开的有效载荷可为但不限于肽、多肽、蛋白质、抗体、RNAi剂等。

[1384] 在一些实施方案中,编码有效载荷的核苷酸序列可包含编码与非编码核酸序列的组合。在一些实施方案中,编码有效载荷的核苷酸序列可编码编码或非编码RNA。

[1385] 在一些实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)包含编码有效载荷的核酸。在一些实施方案中,所编码的有效载荷包括治疗性蛋白质、抗体、酶、基因组编辑系统的一种或多种组分和/或RNAi剂(例如,dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA)。在一些实施方案中,所编码的有效载荷调节(例如,增加或减少)基因、mRNA、蛋白质或其组合例如在细胞或组织中的存在、水平和/或活性。

[1386] 多肽

[1387] 在一些实施方案中,包含本文所述的AAV衣壳多肽(例如,AAV衣壳变体)的AAV颗粒的所编码的有效载荷包含多肽、蛋白质或肽,例如本文所述的多肽、蛋白质或肽。编码有效载荷的核酸可编码任何已知基因的产物和/或其重组形式。在一些实施方案中,编码有效载荷的核酸可编码载脂蛋白E(APOE)(诸如但不限于ApoE2、ApoE3和/或ApoE4)的至少一个等位基因。在一个实施方案中,编码有效载荷的核酸编码ApoE2(cys112、cys158)蛋白或其片

段或变体。在一个实施方案中,编码有效载荷的核酸编码ApoE3(cys112、arg158)蛋白或其片段或变体。在一个实施方案中,编码有效载荷的核酸编码ApoE4(arg112、arg158)。作为另一非限制性实例,所编码的有效载荷包含芳族L-氨基酸脱羧酶(AADC)蛋白。作为另一非限制性实例,所编码的有效载荷包含抗体或其片段。作为另一非限制性实例,所编码的有效载荷包含人类运动神经元存活蛋白(SMN)1或SMN2蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷区域包含葡糖脑苷脂酶(GBA1)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含颗粒体蛋白前体或前颗粒体蛋白(GRN)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含天冬氨酸酰化酶(ASPA)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含三肽基肽酶I(CLN2)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含β-半乳糖苷酶(GLB1)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含N-磺基葡糖胺磺基水解酶(SGSH)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含N-乙酰基-α-氨基葡糖苷酶(NAGLU)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含艾杜糖2-硫酸酯酶(IDS)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含细胞内胆固醇转运蛋白(NPC1)蛋白或其片段或变体。作为另一非限制性实例,所编码的有效载荷包含巨轴索神经蛋白(GAN)蛋白或其片段或变体。编码本文所述的多肽的AAV病毒基因组可用于人类疾病、病毒、感染兽医应用以及各种体内和体外设定的领域中。

[1388] 由本文所述的病毒基因组编码的有效载荷多肽的氨基酸序列可被转译为完整多肽、多个多肽或多肽片段,其可独立地由一个或多个核酸、核酸片段或前述任一个的变体编码。

[1389] 抗体和抗体结合片段

[1390] 在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒的所编码的有效载荷包含抗体或抗体结合片段。在一些实施方案中,所述抗体可为完整抗体、片段或其任何功能变体。作为非限制性实例,抗体可为原生抗体(例如,具有两条重链和两条轻链)、重链可变区、轻链可变区、重链恒定区、轻链恒定区、Fab、Fab'、$F(ab')_2$、Fv或scFv片段、双功能抗体、线性抗体、单链抗体、多特异性抗体、内抗体、一个或多个重链互补决定区(CDR)、一个或多个轻链CDR、双特异性抗体、单克隆抗体、多克隆抗体、人源化抗体、抗体模拟物、抗体变体、小型化抗体、单抗体、大抗体和/或嵌合抗原受体。所编码的抗体或抗体结合片段可用于治疗神经疾病、神经退行性病症、肌肉疾病、神经肌肉病症、神经肿瘤性病症或与中枢和/或外周神经系统相关的任何病症。

[1391] 在一些实施方案中,所述AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)的病毒基因组可包含已被工程改造以实现或增加抗体或其抗体结合片段的表达的核酸。

[1392] 在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒的有效载荷的编码抗体包含至少一个免疫球蛋白可变结构域序列。抗体可包括例如全长、成熟抗体和抗体的抗原结合片段。例如,抗体可包括重(H)链可变结构域序列(VH)和轻(L)链可变结构域序列(VL)。在另一实例中,抗体包括两个重(H)链可变结构域序列和两个轻(L)链可变结构域序列,由此形成两个抗原结合位点,诸如Fab、Fab'、$F(ab')_2$、Fc、Fd、Fd'、Fv、单链抗体(例如,scFv)、单一可变结构域抗体、双功能抗体(Dab)(二价和双特异性)和嵌合(例如,人源化)抗体,其可通过修饰完整抗体或使用重组DNA技术从头合成的那些而产生。这些功能性抗体片

段(例如,抗体结合片段)保留与其各别抗原或受体选择性结合的能力。

[1393] 在一些实施方案中,抗体结合片段包含完整抗体或其重组变体的至少一部分,且是指足以赋予所述抗体片段的识别及与靶标(诸如抗原)的特异性结合的抗原结合结构域,例如完整抗体的抗原决定可变区。抗原结合片段的实例包括:(i)Fab片段,由VL、VH、CL和CH1结构域组成的单价片段;(ii)F(ab')2片段,包含在铰链区通过二硫桥连接的两个Fab片段的二价片段;(iii)Fd片段,由VH和CH1结构域组成;(iv)Fv片段,由抗体单臂的VL和VH结构域组成,(v)双功能抗体(dAb)片段,其由VH结构域组成;(vi)骆驼科动物或骆驼化可变结构域;(vii)单链Fv(scFv),参见,例如Bird等人(1988)Science242:423-426;及Huston等人(1988)Proc.Natl.Acad.Sci.USA 85:5879-5883);和(viii)单一结构域抗体。使用本领域技术人员已知的常规技术获得这些抗体片段,且以与完整抗体相同的方式筛选所述片段的效用。抗体片段也可并入单一结构域抗体、大抗体、微型抗体、纳米抗体、内抗体、双功能抗体、三功能抗体、四功能抗体、v-NAR和双-scFv(参见,例如Hollinger和Hudson,Nature Biotechnology23:1126-1136,2005)中。

[1394] 在一些实施方案中,本文所述的AAV颗粒的有效载荷的编码抗体包含多特异性抗体,例如,其包含多个免疫球蛋白可变结构域序列,其中所述多个中的第一免疫球蛋白可变结构域序列对第一表位具有结合特异性且所述多个中的第二免疫球蛋白可变结构域序列对第二表位具有结合特异性。在一些实施方案中,第一和第二表位在相同抗原,例如相同蛋白质(或多聚体蛋白质的亚单位)上。在一些实施方案中,第一表位与第二表位重叠。在一些实施方案中,第一表位与第二表位不重叠。在一些实施方案中,第一和第二表位在不同抗原,例如不同蛋白质(或多聚体蛋白质的不同亚单位)上。在一些实施方案中,多特异性抗体包含第三、第四或第五免疫球蛋白可变结构域。在一些实施方案中,多特异性抗体为双特异性抗体、三特异性抗体或四特异性抗体。

[1395] 在一些实施方案中,本文所述的AAV颗粒的有效载荷的所编码的多特异性抗体为所编码的双特异性抗体。双特异性抗体对不超过两种抗原具有特异性。双特异性抗体的特征在于第一免疫球蛋白可变结构域序列对第一表位具有结合特异性且第二免疫球蛋白可变结构域序列对第二表位具有结合特异性。在一些实施方案中,第一和第二表位在相同抗原,例如相同蛋白质(或多聚体蛋白质的亚单位)上。在一些实施方案中,第一表位与第二表位重叠。在一些实施方案中,第一表位与第二表位不重叠。在一些实施方案中,第一和第二表位在不同抗原,例如不同蛋白质(或多聚体蛋白质的不同亚单位)上。

[1396] 由本文所述的AAV颗粒的病毒基因组编码的抗体或抗体结合片段可为但不限于与β-淀粉样蛋白、APOE、tau、SOD1、TDP-43、亨廷顿蛋白和/或突触核蛋白结合的抗体或抗体片段。在一些实施方案中,所编码的有效载荷包含与神经肿瘤学相关靶标,例如HER2、EGFR(例如,EGFRvIII)结合的抗体或抗体片段。在一些实施方案中,所编码的有效载荷包含与HER2/neu结合的抗体。在一些实施方案中,所编码的有效载荷包含与β-淀粉样蛋白结合的抗体。在一些实施方案中,所编码的有效载荷包含与tau结合的抗体。

[1397] 基因编辑过程

[1398] 在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒的所编码的有效载荷包含基因编辑系统或其一种或多种组分。在一些实施方案中,基因编辑系统包含编码具有酶活性的蛋白质的核酸序列,以(i)选择性地诱导DNA或RNA序列中的双链或单链断裂,或

(ii)在DNA或RNA中不存在双链或单链断裂的情况下,取代、插入或缺失DNA或RNA序列中的特定碱基或碱基集合。在一些实施方案中,基因编辑系统包括但不限于CRISPR-Cas系统(包括不同Cas或Cas相关核酸酶)、锌指核酸酶、大范围核酸酶、TALEN或碱基编辑器。在一些实施方案中,基因编辑系统包含转基因的染色体整合,例如,在不存在外源核酸酶或酶实体的情况下由微小病毒载体引入。

[1399] RNAi剂

[1400] 在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒的所编码的有效载荷包含RNAi剂,例如本文所述的RNAi剂。在一些实施方案中,包含本文所述的AAV衣壳变体的AAV颗粒的病毒基因组的所编码的有效载荷包含RNAi剂,所述RNAi诸如但不限于dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、stRNA、lncRNA、piRNA或snoRNA。在一些实施方案中,所编码的有效载荷包含用于抑制SOD1、MAPT、APOE、HTT、C9ORF72、TDP-43、APP、BACE、SNCA、ATXN1、ATXN3、ATXN7、SCN1A-SCN5A或SCN8A-SCN11A基因、蛋白质和/或mRNA的表达的RNAi剂。在一些实施方案中,由本文所述的病毒基因组编码的RNAi剂抑制SOD1、MAPT、APOE、HTT、C9ORF72、TDP-43、APP、BACE、SNCA、ATXN1、ATXN3、ATXN7、SCN1A-SCN5A或SCN8A-SCN11A。

[1401] 包含本文所述的AAV衣壳变体的AAV颗粒可包含编码RNAi剂的病毒基因组,所述RNAi剂靶向基因的mRNA以调节(例如,干扰)基因表达和/或蛋白质产生。

[1402] 在一些实施方案中,RNAi剂可在基因的核苷酸序列内的单核苷酸多态性(SNP)或变体的位置处靶向所述基因。

[1403] RNAi剂可为siRNA双链体,其中所述siRNA双链体含有杂交在一起形成双链体结构的反义链(引导链)和有义链(过客链),其中反义链与靶向基因的核酸序列互补,且其中有义链与靶向基因的核酸序列同源。在一些方面中,反义链的5'端具有5'磷酸酯基且有义链的3'端含有3'羟基。在其他方面中,在每个链的3'端不存在、存在一个或2个核苷酸悬垂物。

[1404] 靶向相关基因的siRNA双链体的每个链可为约19至25、19至24或19至21个核苷酸的长度,优选地约19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸或25个核苷酸的长度。

[1405] 在一个实施方案中,siRNA或dsRNA包括至少两个彼此互补的序列。dsRNA包括具有第一序列的有义链和具有第二序列的反义链。反义链包括与编码靶基因的mRNA的至少一部分大体上互补的核苷酸序列,且互补区域为30个核苷酸或更少且至少15个核苷酸的长度。一般来说,dsRNA为19至25、19至24或19至21个核苷酸的长度。在一些实施方案中,dsRNA为约15至约25个核苷酸的长度,且在其他实施方案中,dsRNA为约25至约30个核苷酸的长度。在一些实施方案中,dsRNA为约15个核苷酸的长度、16个核苷酸的长度、17个核苷酸的长度、18个核苷酸的长度、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸的长度、26个核苷酸的长度、27个核苷酸的长度、28个核苷酸的长度、29个核苷酸的长度或30个核苷酸的长度。

[1406] 在一些实施方案中,所编码的RNAi剂为siRNA。

[1407] 在一些实施方案中,RNAi剂(例如,本文所述的RNAi剂)抑制基因、mRNA和/或蛋白质的表达达至少10%、至少20%、至少25%、至少30%、至少35%或至少40%或更多,诸如当通过本领域中已知的方法分析时。在一些实施方案中,RNAi剂抑制基因、mRNA和蛋白质的表达达50-100%,例如达30%、40%、50%、60%、70%、80%、85%、90%、95%和100%。

[1408] 在一些实施方案中,将本文所述的包含编码靶向相关基因的RNAi剂的病毒基因组的AAV颗粒施用于需要治疗和/或改善疾病(例如,与中枢或外周神经系统相关的任何疾病的神经病症)的受试者。

[1409] siRNA的设计

[1410] 本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)可包含编码siRNA分子(例如,siRNA双链体或所编码的dsRNA)的病毒基因组,所述siRNA分子靶向目标基因且抑制靶基因表达、mRNA表达及蛋白质产生。在一些方面中,siRNA分子被设计且用于敲除细胞中的靶基因变体,例如在神经疾病中鉴定的转录本。在一些方面中,siRNA分子被设计且用于敲低细胞中的靶基因变体。

[1411] 本领域中已提出一些用于设计siRNA(用于插入至本文所述的AAV颗粒的病毒基因组中)的指南。这些指南一般推荐生成19个核苷酸的双链区域、对称的2-3个核苷酸的3'悬垂物、5-磷酸酯基和3-羟基,靶向基因中待沉默的区域。可管理siRNA序列偏好的其他规则包括但不限于(i)在反义链的5′端的A/U;(ii)在有义链的5′端的G/C;(iii)在反义链的5′末端三分之一处的至少五个A/U残基;及(iv)不存在超过9个核苷酸的长度的任何GC延伸段。根据此类考虑,连同靶基因的特定序列,可容易地设计抑制哺乳动物靶基因表达所必需的高效siRNA分子。

[1412] 在一个实施方案中,有义和/或反义链是基于欧洲专利公布号EP1752536中所概述的方法和规则进行设计,其内容以引用方式整体并入本文中。作为非限制性实例,序列的3'端碱基为腺嘌呤、胸腺嘧啶或尿嘧啶。作为非限制性实例,序列的5'端碱基为鸟嘌呤或胞嘧啶。作为非限制性实例,3'端序列包含富含腺嘌呤、胸腺嘧啶和尿嘧啶的一个或多个碱基的七个碱基。

[1413] 在一个实施方案中,siRNA分子包含有义链和互补反义链,其中两个链杂交在一起以形成双链体结构。反义链与靶mRNA序列具有足够互补性以指导靶标特异性RNAi,例如,siRNA分子具有足以通过RNAi机制或过程触发靶mRNA的破坏的序列。

[1414] 在一些实施方案中,反义链与靶mRNA序列具有100%互补性。反义链可与靶mRNA序列的任何部分互补。有义序列的同一性或反义序列的同源性均不需要与靶标100%互补。

[1415] 在其他实施方案中,反义链与靶mRNA序列包含至少一个错配。作为非限制性实例,反义链与靶mRNA序列具有至少50-90%、50-95%、50-99%、60-70%、60-80%、60-90%、60-95%、60-99%、70-80%、70-90%、70-95%、70-99%、80-90%、80-95%、80-99%、90-95%、90-99%或95-99%互补。

[1416] siRNA分子可具有约10-50个或更多核苷酸的长度,例如,每个链包含10-50个核苷酸(或核苷酸类似物)。优选地,siRNA分子具有每个链中约15-30个,例如15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个核苷酸的长度,其中一个链与靶区域充分互补。在一个实施方案中,siRNA分子具有约19至25、19至24或19至21个核苷酸的长度。

[1417] 在一些实施方案中,siRNA分子可为合成RNA双链体,其包含约19个核苷酸至约25个核苷酸以及在3'端的两个悬垂核苷酸。

[1418] siRNA分子可包含反义序列和有义序列,或其片段或变体。作为非限制性实例,反义序列与有义序列具有至少50-90%、50-95%、50-99%、60-70%、60-80%、60-90%、60-95%、60-99%、70-80%、70-90%、70-95%、70-99%、80-90%、80-95%、80-99%、90-95%、

90-99%或95-99%互补。

[1419] 有义与反义序列可在其长度的大部分上完全互补。在其他实施方案中,有义序列与反义序列可独立地在这些链的长度的至少50、60、70、80、85、90、95或99%上至少70、80、90、95或99%互补。

[1420] 在一些实施方案中,siRNA双链体的有义链与反义链通过短间隔序列连接,导致称为短发夹RNA(shRNA)的茎环结构的表达。发夹由Dicer识别且裂解,因此产生成熟siRNA分子。

[1421] 在一些实施方案中,一旦被设计,siRNA分子以及相关间隔子和/或侧接区域即可由本文所述的AAV颗粒的病毒基因组编码,以递送至细胞。

[1422] 分子支架

[1423] 在一些实施方案中,siRNA分子可在调节性多核苷酸中编码,调节性多核苷酸还包含分子支架。

[1424] 在一些实施方案中,包含有效载荷(例如,siRNA、miRNA或本文所述的其他RNAi剂)的调节性多核苷酸包括分子支架,分子支架包含5'侧接序列、环区和/或3'侧接区域。在一些实施方案中,5'或3'侧接区域可具有任何长度且可为野生型微小RNA序列或其一部分,或可为完全人工的。3'侧接序列在大小和来源上可反映5'侧接序列。任一侧接序列均可能不存在。在一个实施方案中,5'和3'侧接序列均不存在。3'侧接序列可任选地含有一个或多个CNNC模体,其中"N"表示任何核苷酸。在一些实施方案中,环包含至少一个UGUG模体。在一些实施方案中,UGUG模体位于环的5'端。在一些实施方案中,5'和3'侧接序列为相同序列。在一些实施方案中,当彼此比对时,这些序列相差2%、3%、4%、5%、10%、20%或超过30%。

[1425] 在一些实施方案中,调节性多核苷酸包含茎环结构。在一些实施方案中,调节性多核苷酸依5'至3'次序包含:5'侧接序列、引导链序列、环区、过客链序列和3'侧接序列。在一些实施方案中,调节性多核苷酸依5'至3'次序包含:5'侧接序列、过客链序列、环区、引导链序列和3'侧接序列。

[1426] 在一个实施方案中,分子支架包含双功能靶向调节性多核苷酸。在一个实施方案中,分子支架可包含本领域中已知的一个或多个接头。接头可将区域或一个分子支架与另一者分开。作为非限制性实例,分子支架可为多顺反子的。

[1427] 在一个实施方案中,使用至少一种以下特性来设计调节性多核苷酸:环变体、种子错配/凸起/摆动变体、茎错配、环变体和基础茎错配变体、种子错配和基础茎错配变体、茎错配和基础茎错配变体、种子摆动和基础茎摆动变体或者茎序列变体。

[1428] AAV产生

[1429] 本文所公开的病毒产生描述用于产生AAV颗粒(对靶组织具有增强、改善和/或增加的趋向性)的过程和方法,例如包含AAV衣壳变体的AAV颗粒,其可用于接触靶细胞以递送有效载荷。

[1430] 在一些实施方案中,本文揭示一种制备本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的方法,所述方法包括:(i)提供包含本文所述的病毒基因组的宿主细胞,及(ii)在适合将所述病毒基因组封入AAV衣壳变体(例如,本文所述的AAV衣壳变体(例如,表3、表4或表5中所列的AAV衣壳变体))中的条件下培育宿主细胞,由此制备所述AAV颗粒。在一些实施方案中,所述方法包括在步骤(i)之前,将包含病毒基因组的第一核酸引入细胞

中。在一些实施方案中,宿主细胞包含编码AAV衣壳变体的第二核酸。在一些实施方案中,在第一核酸分子之前、同时或之后将第二核酸引入宿主细胞中。在一些实施方案中,本文所述的AAV颗粒为分离的AAV颗粒。在一些实施方案中,本文所述的AAV颗粒为重组AAV颗粒。

[1431] 本领域中已知的任何方法均可用于制备AAV颗粒。在一些实施方案中,AAV颗粒在哺乳动物细胞(例如,HEK293)中产生。在另一实施方案中,AAV颗粒在昆虫细胞(例如,Sf9)中产生。

[1432] 制备AAV颗粒的方法为本领域中熟知的且描述于例如美国专利号US6204059、US5756283、US6258595、US6261551、US6270996、US6281010、US6365394、US6475769、US6482634、US6485966、US6943019、US6953690、US7022519、US7238526、US7291498和US7491508、US5064764、US6194191、US6566118、US8137948;或国际公布号WO1996039530、WO1998010088、WO1999014354、WO1999015685、WO1999047691、WO2000055342、WO2000075353和WO2001023597;Methods In Molecular Biology,Richard编,Humana Press,NJ(1995);O'Reilly等人,Baculovirus Expression Vectors,ALaboratory Manual,Oxford Univ.Press(1994);Samulski等人,J.Vir.63:3822-8(1989);Kajigaya等人,Proc.Nat'l.Acad.Sci.USA 88:4646-50(1991);Ruffing等人,J.Vir.66:6922-30(1992);Kimbauer等人,Vir.,219:37-44(1996);Zhao等人,Vir.272:382-93(2000)中;其各自的内容均以引用方式整体并入本文中。在一些实施方案中,AAV颗粒使用国际专利公布WO2015191508中所述的方法制得,该专利的内容以引用方式整体并入本文中。

[1433] 治疗应用

[1434] 本公开提供一种用于治疗受试者(包括人受试者)的疾病、病症和/或疾患的方法,所述方法包括向受试者施用本文所述的AAV颗粒,例如包含AAV衣壳变体(例如,本文所述的AAV衣壳变体)的AAV颗粒,或向受试者施用本文所述的任何所述组合物,包括药物组合物。

[1435] 在一些实施方案中,将本公开的AAV颗粒(例如,包含本文所述的AAV衣壳变体,例如AAV5衣壳变体的AAV颗粒)预防性地施用于受试者,以预防疾病发作。在另一实施方案中,施用本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)以治疗疾病或其症状(例如,减轻其效应)。在又一实施方案中,施用本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)以治愈(消除)疾病。在另一实施方案中,施用本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)以预防或减缓疾病进展。在又一实施方案中,使用本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)来逆转疾病的有害效应。可通过本领域中已知的标准方法来确定或监测疾病状态和/或进展。

[1436] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善遗传病症,例如,常染色体显性遗传病症、常染色体隐性病症、X-连锁显性遗传病症、X-连锁隐性遗传病症或Y-连锁遗传病症。在一些实施方案中,遗传病症为单基因病症或多基因病症。在一些实施方案中,遗传病症(例如,单基因病症)的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。

[1437] 在一些实施方案中,本文提供用于治疗受试者的神经病症和/或神经退行性病症的方法,所述方法包括向受试者施用有效量的本文所述的药物组合物或AAV颗粒(例如,多个颗粒),药物组合物或AAV颗粒包含本文所述的AAV衣壳变体。在一些实施方案中,神经病症和/或神经退行性病症的治疗包括预防该神经病症和/或神经病症。

[1438] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善神经疾病和/或病症。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善tau蛋白病变。

[1439] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)用于治疗、预防、缓解或改善阿尔茨海默氏病。在一些实施方案中,阿尔茨海默氏病的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含ApoE2蛋白、ApoE4蛋白、ApoE3蛋白、BDNF蛋白、CYP46A1蛋白、Klotho蛋白、分形趋化因子(FKN)蛋白、脑啡肽酶蛋白(NEP)、CD74蛋白、caveolin-1或其组合或变体。在一些实施方案中,阿尔茨海默氏病的治疗包括使用本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)来减少tau基因和/或蛋白质、突触核蛋白基因和/或蛋白质或其组合或变体的表达。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含与tau或突触核蛋白结合的抗体、用于抑制tau或突触核蛋白的RNAi剂、用于改变tau或突触核蛋白表达的基因编辑系统(例如,CRISPR-Cas系统)或其组合。

[1440] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)用于治疗、预防、缓解或改善额颞叶痴呆。在一些实施方案中,额颞叶痴呆的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含前颗粒体蛋白或其变体。

[1441] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善帕金森氏病。在一些实施方案中,帕金森氏病的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含AADC蛋白、GAD蛋白、GDNF蛋白、TH-GCH1蛋白、GBA蛋白、AIMP2-DX2蛋白或其组合或变体。在一些实施方案中,帕金森氏病的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因敲低疗法或基因编辑疗法(例如,敲除、抑制或校正)。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含调节剂,例如RNAi剂或CRISPR-Cas系统,用于改变α-突触核蛋白基因、mRNA和/或蛋白质或其变体的表达。

[1442] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善AADC缺乏。在一些实施方案中,AADC缺乏的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含AADC蛋白或其变体。

[1443] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善肌萎缩性侧索硬化(ALS)。在一些实施方案中,ALS的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含TDP-43蛋白、UPF1蛋白、C9orf72蛋白、CCNF蛋白、HSF1蛋白、H因子蛋白、NGF蛋白、ADAR2蛋白、GDNF蛋白、VEGF蛋白、HGF蛋白、NRTN蛋白、AIMP2-DX2蛋白或其组合或变体。在一些实施方案中,ALS的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因

敲低疗法或基因编辑疗法(例如,敲除、抑制或校正)。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含调节剂,例如RNAi剂或CRISPR-Cas系统,用于改变SOD1或C9ORF72基因、mRNA和/或蛋白质或其组合或变体的表达。

[1444] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善亨廷顿氏病。在一些实施方案中,ALS的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因敲低(例如,敲除)疗法或基因编辑疗法(例如,敲除、抑制或校正)。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含调节剂,例如RNAi剂或CRISPR-Cas系统,用于改变HTT基因、mRNA和/或蛋白质或其变体的表达。

[1445] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善脊髓性肌萎缩。在一些实施方案中,脊髓性肌萎缩的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含SMN1蛋白、SMN2蛋白或其组合或变体。

[1446] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善多系统萎缩。在一些实施方案中,多系统萎缩的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。

[1447] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善戈谢病(GD)(例如,1型GD、2型GD或3型GD)。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善与GBA突变相关的帕金森氏病。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善路易体痴呆(DLB)。

[1448] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善脑白质营养不良(例如,亚历山大病、常染色体显性脑白质营养不良伴自主神经疾病(ADLD)、肾上腺脑白质营养不良(ALD)、卡纳万病、脑腱黄瘤病(CTX)、异染性脑白质营养不良(MLD)、佩-梅二氏病或雷夫叙姆病)。在一些实施方案中,MLD的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含ARSA蛋白或其变体。在一些实施方案中,ALD的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含ABCD-1蛋白或其变体。

[1449] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善巨脑性脑白质病(MLC)。在一些实施方案中,MLC的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含MLC1蛋白或其变体。

[1450] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Krabbe疾病。在一些实施方案中,Krabbe疾病的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实

施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含GALC蛋白或其变体。

[1451] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善粘多糖贮积症(例如,I型(MPS I)、II型(MPS II)、IIIA型(MPS IIIA)、IIIB型(MPS IIIB)或IIIC型(MPS IIIC))。在一些实施方案中,粘多糖贮积症的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法或基因编辑疗法(例如,增强或校正)。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码或校正的有效载荷包含IDUA蛋白、IDS蛋白、SGSH蛋白、NAGLU蛋白、HGSNAT蛋白或其组合或变体。

[1452] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Batten/NCL。在一些实施方案中,Batten/NCL的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含CLN1蛋白、CLN2蛋白、CLN3蛋白、CLN5蛋白、CLN6蛋白、CLN7蛋白、CLN8蛋白或其组合或变体。

[1453] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Rett综合征。在一些实施方案中,Rett综合征的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含MeCP2蛋白或其变体。

[1454] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善天使综合征(Angelman Syndrome)。在一些实施方案中,天使综合征的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含UBE3A蛋白或其变体。

[1455] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善脆性X综合征。在一些实施方案中,脆性X综合征的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含Reelin蛋白、DgkK蛋白、FMR1蛋白或其组合或变体。

[1456] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善卡纳万病。在一些实施方案中,卡纳万病的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含ASPA蛋白或其变体。

[1457] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善神经节苷脂症(例如,GM1神经节苷脂症或GM2神经节苷脂症(例如,Tay Sachs Sandhoff))。在一些实施方案中,神经节苷脂症(例如,GM1神经节苷脂症或GM2神经节苷脂症(例如,Tay Sachs Sandhoff))的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含GLB1蛋白、HEXA蛋白、HEXB蛋白、GM2A蛋白或其组合或变体。

[1458] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善GM3合成酶缺乏。在一些实施方案中,GM3合成酶缺乏的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含ST3GAL5蛋白或其变体。

[1459] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善尼曼匹克病(Niemann-Pick disorder)(例如,尼曼匹克A或尼曼匹克C1(NPC-1))。在一些实施方案中,尼曼匹克病(例如,尼曼匹克A或尼曼匹克C1(NPC-1))的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含ASM蛋白、NPC1蛋白或其变体。

[1460] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善神经鞘瘤(例如,神经瘤)。在一些实施方案中,神经鞘瘤(例如,神经瘤)的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含半胱天冬酶-1蛋白或其变体。

[1461] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善结节性硬化症(例如,1型结节性硬化症或2型结节性硬化症)。在一些实施方案中,结节性硬化症(例如,1型结节性硬化症或2型结节性硬化症)的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含TSC1蛋白、TSC2蛋白或其变体。

[1462] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善CDKL5缺乏。在一些实施方案中,CDKL5缺乏的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含CDKL5蛋白或其变体。

[1463] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Charcot-Marie-Tooth病症(例如,Charcot-Marie-Tooth 1X型(CMT1X)病症、Charcot-Marie-Tooth 2A型(CMT2A)病症或Charcot-Marie-Tooth 4J型(CMT4J)病症)。在一些实施方案中,Charcot-Marie-Tooth病症(例如,Charcot-Marie-Tooth 1X型(CMT1X)病症、Charcot-Marie-Tooth 2A型(CMT2A)病症或Charcot-Marie-Tooth 4J型(CMT4J)病症)的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含GJB1蛋白、MFN2蛋白、FIG4蛋白或其变体。

[1464] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善天冬胺酰葡糖胺尿症(AGU)。在一些实施方案中,AGU的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含AGA蛋白或其变体。

[1465] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Leigh综合征。在一些实施方案中,Leigh综合征的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含SURF1蛋白或其变体。

[1466] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善癫痫。在一些实施方案中,癫痫的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含NPY/Y2蛋白、甘丙肽蛋白、强啡肽蛋白、AIMP2-DX2蛋白、SLC6A1蛋白、SLC13A5蛋白、KCNQ2蛋白或其变体。

[1467] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Dravet综合征。在一些实施方案中,Dravet综合征的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含SCN1a蛋白或其变体。

[1468] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善杜兴氏肌肉萎缩(DMD)。在一些实施方案中,DMD的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法或增强(例如,外显子跳过的校正)或基因编辑疗法(例如,增强或校正)。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码或校正的有效载荷包含抗肌萎缩蛋白基因和/或蛋白质、肌营养相关蛋白基因和/或蛋白质或者GALGT2基因和/或蛋白质或者滤泡抑素基因和/或蛋白质,或其组合或变体。

[1469] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善Pompe疾病。在一些实施方案中,Pompe疾病的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含GAA蛋白或其变体。

[1470] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善肢带肌肉萎缩(LGMD2A)。在一些实施方案中,LGMD2A的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。在一些实施方案中,由包含本文所述的衣壳变体的AAV颗粒编码的有效载荷包含CAPN-3蛋白、DYSF蛋白、SGCG蛋白、SGCA蛋白、SGCB蛋白、FKRP蛋白、ANO5蛋白或其组合或变体。

[1471] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善慢性或神经病理性疼痛。

[1472] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善与中枢神经系统相关的疾病。

[1473] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善与外周神经系统相关的疾病。

[1474] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善受试者的神经肿瘤性病症。在一些实施方案中,神经肿瘤性病症的

治疗包括预防该神经肿瘤性病症。在一些实施方案中,神经肿瘤性病症包含原发性CNS起源(例如,CNS细胞、组织或区域)的癌症,或CNS细胞、组织或区域中的转移性癌症。原发性CNS癌症的实例可能为神经胶质瘤(其可包括胶质母细胞瘤(也称为多形性胶质母细胞瘤)、星形细胞瘤、寡树突神经胶质瘤和室管膜瘤及混合性神经胶质瘤)、脑膜瘤、神经管胚细胞瘤、神经瘤和原发性CNS淋巴瘤(在脑、脊髓或脑膜中)等等。转移性癌症的实例包括起源于另一组织或器官的那些,例如乳腺癌、肺癌、淋巴瘤、白血病、黑色素瘤(皮肤癌)、结肠癌、肾癌、前列腺癌或转移至脑的其他类型。

[1475] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善与HER2表达相关的疾病(例如,与HER2过表达相关的疾病)。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善HER2阳性癌症。在一些实施方案中,HER2阳性癌症为HER2阳性实体肿瘤。另外或替代地,HER2阳性癌症可为局部晚期或转移性HER2阳性癌症。在一些情况下,HER2阳性癌症为HER2阳性乳腺癌或HER2阳性胃癌。在一些实施方案中,HER2阳性癌症选自由HER2阳性胃食道结合部癌、HER2阳性结肠直肠癌、HER2阳性肺癌(例如,HER2阳性非小细胞肺癌)、HER2阳性胰脏癌、HER2阳性结肠直肠癌、HER2阳性膀胱癌、HER2阳性唾液腺管癌、HER2阳性卵巢癌(例如,HER2阳性上皮卵巢癌)或HER2阳性子宫内膜癌组成的组。在一些情况下,HER2阳性癌症为前列腺癌。在一些实施方案中,HER2阳性癌症已转移到中枢神经系统(CNS)。在一些情况下,转移的HER2癌症已形成CNS赘瘤。

[1476] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善受试者的肌肉病症和/或神经肌肉病症。在一些实施方案中,肌肉病症和/或神经肌肉病症的治疗包括预防该肌肉病症和/或神经肌肉病症。

[1477] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善心脏病(cardiac disease/heart disease)和/或改善(例如,增强)受试者的心脏功能的方法。在一些实施方案中,心脏病为心肌病(例如,致心律失常性右心室心肌病、扩张型心肌病或肥厚性心肌病)、充血性心脏衰竭、心动过速(例如,儿茶酚胺激导性多形性室性心动过速)、缺血性心脏病和/或心肌梗塞。在一些实施方案中,心脏病为与LAMP2B、MYBPC3、TNNI3、LMNA、BAG3、DWORF、PKP2、Cx43、TAZ、CASQ2、SERCA2a、I-1c、S100A1和/或ARC、S100A1、ASCL1、miR133、Mydelta3、Sav或其组合或变体的表达(例如,异常表达)相关的疾病。在一些实施方案中,本文所述的心脏病症的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。

[1478] 在一些实施方案中,心脏病为遗传病症(例如,常染色体显性遗传病症、常染色体隐性病症或X-连锁隐性遗传病症)。在一些实施方案中,心肌病为遗传病症,例如与选自TTN、LMNA、MYH7、MYH6、SCN5A、TNNT2、RBM20、TNNI3、MYL2、MYL3、PKP2、DSP、DSG2、DSC2、JUP或其组合的基因的异常(例如,突变、插入、重排和/或缺失)相关的遗传病症。在一些实施方案中,心脏病症为扩张型心肌病,例如与选自TTN、LMNA、MIH7、BAG3、MIPN、TNNT2、SCN5A、RBN20、TNPO、LAMA4、VCL、LDB3、TCAP、PSEN1/2、ACTN2、CRYAB、TPM1、ABCC9、ACTC1、PDLIM3、ILK、TNNC1、TNNI3、PLN、DES、SGCD、CSRP3、MIH6、EYA4、ANKRD1、DMD、GATAD1、TAZ/G4.5或其组合的基因的异常(例如,突变、插入、重排和/或缺失)相关的扩张型心肌病。在一些实施方案中,心脏病症为肥厚性心肌病,例如与选自MYH7、TNNT2、TNNI3、TPM1、MYL2、MYL3、ACTC1、

CSRP3、TTN、ACTN2、MYH6、TCAP、TNNC1或其组合的基因的异常(例如,突变、插入、重排和/或缺失)相关的肥厚性心肌病。在一些实施方案中,心脏病症为致心律失常性心室心肌病,例如与选自PKP2、DSG2、DSP、RYR2、DSC2、TGFB3、TMEM43、DES、TTN、LMNA或其组合的基因的异常(例如,突变、插入、重排和/或缺失)相关的致心律失常性心室心肌病。

[1479] 在一些实施方案中,将本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)施用于具有至少一种本文所述的疾病或症状的受试者。在一些实施方案中,将本公开的AAV颗粒施用于患有或被诊断患有本文所述的疾病或病症的受试者。

[1480] 任何神经疾病或病症、神经退行性病症、肌肉病症、神经肌肉病症和/或神经肿瘤性病症均可用本公开的AAV颗粒或其药物组合物治疗。

[1481] 药物组合物和制剂

[1482] 根据本公开,可将包含本文所述的AAV衣壳变体的AAV颗粒制备成药物组合物。在一些实施方案中,药物组合物包含至少一种活性成分。在一些实施方案中,药物组合物包含药学上可接受的赋形剂。

[1483] 在一些实施方案中,可使用赋形剂来配制本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)以:(1)提高稳定性;(2)增加细胞转染或转导;(3)允许有效载荷的持续或延迟表达;(4)改变生物分布(例如,将病毒颗粒靶向特定组织或细胞类型);(5)增加经编码蛋白质的转译;(6)改变经编码蛋白质的释放型态;和/或(7)允许有效载荷的可调节表达。本公开的制剂可包括但不限于生理食盐水、脂质体、脂质纳米颗粒、聚合物、肽、蛋白质、病毒载体转染的细胞(例如,用于转移或移植到受试者中)及其组合。

[1484] 在一些实施方案中,根据本公开的药物组合物中的活性成分(例如,包含本文所述的AAV衣壳变体的AAV颗粒)、药学上可接受的赋形剂和/或任何额外成分的相对量可根据所治疗受试者的身份、体格和/或状况且进一步根据待用于施用组合物的途径而变化。例如,组合物可包含0.1%与99%(w/w)之间的活性成分。举例来说,组合物可包含0.1%与100%之间,例如.5与50%之间、1-30%之间、5-80%之间、至少80%(w/w)的活性成分。

[1485] 在一些实施方案中,包含本文所述的AAV颗粒的药物组合物可包含AAV衣壳变体和编码有效载荷(例如,本文所述的有效载荷)的病毒基因组,具有或不具有药学上可接受的赋形剂。

[1486] 在一些实施方案中,本公开还提供一种适合向受试者(例如,人)施用的药物组合物。在一些实施方案中,将所述药物组合物向受试者(例如,人)施用。

[1487] 施用

[1488] 在一些实施方案中,本文所公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过递送途径(例如,局部递送途径或全身递送途径)施用于受试者。

[1489] 在一些实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可经由使其能够穿过血脑障壁、血管障壁或其他上皮障壁的途径施用。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可以任何合适形式施用,无论作为液体溶液或悬浮液,抑或作为适合液体溶液或液体溶液中的悬浮液的固体形式。在一些实施方案中,AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可与任何适当且药学上可接受的赋形剂一起配制。

[1490] 在一些实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)被肌

肉内、静脉内、脑内、鞘内、脑室内、经由实质内施用或经由大池内注射(ICM)施用。

[1491] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可经由单一途径施用递送至受试者。在一些实施方案中,本公开的AAV颗粒可经由多位点施用途径递送至受试者。在一些实施方案中,可在2、3、4、5个或超过5个位点处对受试者进行施用。

[1492] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)经由推注施用。在一些实施方案中,本公开的AAV颗粒经由数分钟、数小时或数天的时期内的持续递送来施用。在一些实施方案中,输注速率可根据受试者、分布、配方和/或另一递送参数而改变。在一些实施方案中,本公开的AAV颗粒使用控制释放来施用。在一些实施方案中,本公开的AAV颗粒使用持续释放(例如,符合一段特定时期内的释放速率的释放型态)来施用。

[1493] 在一些实施方案中,AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过超过一种施用途径来递送。作为组合施用的非限制性实例,AAV颗粒可通过鞘内和脑室内,或通过静脉内和实质内施用来递送。

[1494] 静脉内施用

[1495] 在一些实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)可通过全身施用施用于受试者。在一些实施方案中,全身施用为静脉内施用。在另一实施方案中,全身施用为动脉内施用。在一些实施方案中,本公开的AAV颗粒可通过静脉内施用施用于受试者。在一些实施方案中,静脉内施用可通过皮下递送实现。在一些实施方案中,AAV颗粒经由聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合施用于受试者,例如,如Terstappen等人(Nat Rev Drug Discovery,doi.org/10.1038/s41573-021-00139-y(2021))中所述,其内容以引用方式整体并入本文中。在一些实施方案中,AAV颗粒(例如,本文所述的AAV颗粒)静脉内施用于受试者。在一些实施方案中,所述受试者为人。

[1496] 施用于CNS

[1497] 在一些实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过直接注射到脑中来递送。作为非限制性实例,脑递送可通过海马体内施用实现。在一些实施方案中,本公开的AAV颗粒可通过实质内施用施用于受试者。在一些实施方案中,实质内施用是针对中枢神经系统的组织。在一些实施方案中,本公开的AAV颗粒可通过颅内递送施用于受试者(参见例如美国专利号8119611;其内容以引用方式整体并入本文中)。在一些实施方案中,本文所述的AAV颗粒可通过注射到CSF路径中进行递送。递送至CSF路径的非限制性实例包括鞘内和脑室内施用。在一些实施方案中,本文所述的AAV颗粒可经由大池内(ICM)注射施用。

[1498] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过全身递送递送至脑。作为非限制性实例,全身递送可通过血管内施用实现。作为非限制性实例,全身或血管内施用可为静脉内。

[1499] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过眼内递送途径递送。眼内施用的非限制性实例包括玻璃体内注射。

[1500] 肌肉内注射

[1501] 在一些实施方案中,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过肌肉内施用来递送。不希望受理论束缚,据信在一些实施方案中,肌肉细胞的多核性质为

AAV递送后的基因转导提供优势。在一些实施方案中,肌肉细胞能够表达具有适当转译后修饰的重组蛋白。不希望受理论束缚,据信在一些实施方案中,具有血管结构的肌肉组织的富集允许转移至血流和全身递送。肌肉内施用的实例包括全身(例如,静脉内)、皮下或直接施用于肌肉中。在一些实施方案中,施用超过一次注射。在一些实施方案中,本公开的AAV颗粒可通过肌肉内递送途径来递送。(参见,例如美国专利号6506379;其内容以引用方式整体并入本文中)。肌肉内施用的非限制性实例包括静脉内注射或皮下注射。

[1502] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)施用于受试者且转导受试者的肌肉。作为非限制性实例,通过肌肉内施用来施用AAV颗粒。在一些实施方案中,本公开的AAV颗粒可通过皮下施用施用于受试者。在一些实施方案中,肌肉内施用经由全身递送来实施。在一些实施方案中,肌肉内施用经由静脉内递送来实施。在一些实施方案中,肌肉内施用经由直接注射到肌肉来实施。

[1503] 在一些实施方案中,通过施用(例如,肌肉内施用)来转导肌肉。在一些实施方案中,肌肉内递送包含一个位点处的施用。在一些实施方案中,肌肉内递送包含超过一个位点处的施用。在一些实施方案中,肌肉内递送包含两个、三个、四个或更多个位点处的施用。在一些实施方案中,使肌肉内递送与至少一种其他施用方法组合。

[1504] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过外周注射施用于受试者。外周注射的非限制性实例包括腹膜内、肌肉内、静脉内、结膜或关节注射。本领域中已揭示,AAV颗粒的外周施用可转运至中枢神经系统,例如至运动神经元(例如,美国专利公布号20100240739和20100130594;其各自的内容以引用方式整体并入本文中)。

[1505] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可通过实质内施用施用于受试者。在一些实施方案中,实质内施用是针对肌肉组织。在一些实施方案中,本公开的AAV颗粒如Bright等人2015(Neurobiol Aging.36(2):693-709)中所述进行递送,其内容以引用方式整体并入本文中。在一些实施方案中,将本公开的AAV颗粒施用于受试者的腓肠肌。在一些实施方案中,将本公开的AAV颗粒施用于受试者的股二头肌。在一些实施方案中,将本公开的AAV颗粒施用于胫骨前肌。在一些实施方案中,将本公开的AAV颗粒施用于比目鱼肌。

[1506] 储槽施用

[1507] 如本文所述,在一些实施方案中,将本公开的药物组合物和/或AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)配制成储槽以用于延长释放。一般来说,靶向特定器官或组织以进行施用。

[1508] 在一些实施方案中,本公开的药物组合物和/或AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)在空间上保留于靶组织内或接近靶组织。提供将药物组合物、AAV颗粒提供至哺乳动物受试者的靶组织的方法,其中使靶组织(其包含一个或多个靶细胞)与药物组合物和/或所述AAV颗粒在使得其大体上经保留于靶组织中,例如使得组合物的至少10、20、30、40、50、60、70、80、85、90、95、96、97、98、99、99.9、99.99或大于99.99%保留于靶组织中的条件下接触。在一些实施方案中,通过测量进入靶细胞或多个靶细胞的药物组合物和/或AAV颗粒的量来确定保留。例如,在施用后的一段时期内,施用于受试者的药物组合物和/或AAV颗粒的至少1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、

85％、90％、95％、96％、97％、98％、99％、99.9％、99.99％或大于99.99％存在于细胞内。例如,可使用包含本公开的药物组合物和/或AAV颗粒以及转染试剂的水性组合物对受试者进行肌肉内注射,且通过测量存在于肌肉细胞或多个肌肉细胞中的药物组合物和/或AAV的量来确定保留。

[1509] 在一些实施方案中,本文揭示向受试者的组织提供本公开的药物组合物和/或AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)的方法,其中使组织(包含细胞,例如多个细胞)与药物组合物和/或所述AAV颗粒在使得其大体上保留于组织中的条件下接触。在一些实施方案中,本文所述的药物组合物和/或AAV颗粒包含足量活性成分,使得在至少一个细胞中产生相关效应。在一些实施方案中,药物组合物和/或AAV颗粒一般包含一种或多种细胞渗透剂。在一些实施方案中,本公开提供裸制剂(诸如不含细胞渗透剂或其他剂),具有或不具有药学上可接受的载剂。

[1510] 治疗方法

[1511] 本公开提供用于将本公开的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)引入(例如,递送)至细胞中的方法。在一些实施方案中,所述方法包括向细胞中引入本文所述的AAV颗粒或载体,其量足以调节(例如,增加)靶基因、mRNA和/或蛋白质的产生。在一些实施方案中,所述方法包括向细胞中引入本文所述的AAV颗粒或载体,其量足以调节(例如,减少)靶基因、mRNA和/或蛋白质的表达。在一些实施方案中,细胞可为神经元,诸如但不限于运动神经元、海马体神经元、内嗅神经元、丘脑神经元、皮层神经元、感觉神经元、交感神经元或副交感神经元以及神经胶质细胞(诸如星形胶质细胞、小神经胶质细胞和/或少突胶质细胞)。在其他实施方案中,细胞可为肌肉细胞(例如,横膈膜、股四头肌或心脏(例如,心房或心室)的细胞)或肝细胞。在一些实施方案中,细胞可为心脏细胞(例如,心房细胞或心室细胞)。

[1512] 本公开公开了用于治疗与需要治疗的受试者的蛋白质(例如,靶蛋白)的功能/存在异常(例如,不足或增加)相关的神经疾病/病症或神经退行性病症、肌肉或神经肌肉病症或神经肿瘤性病症的方法。

[1513] 在一些实施方案中,所述方法包括向所述受试者施用治疗有效量的包含本公开的AAV颗粒的组合物。作为非限制性实例,AAV颗粒可增加靶基因表达,增加靶蛋白产生,且因此减少受试者的神经疾病的一种或多种症状,使得所述受试者受到治疗性治疗。

[1514] 在其他实施方案中,所述方法包括向所述受试者施用治疗有效量的包含AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)的组合物,所述AAV颗粒包含具有编码一种或多种siRNA分子的核酸序列的病毒基因组。作为非限制性实例,siRNA分子可使靶基因表达沉默,抑制靶蛋白产生,且减轻受试者的神经疾病的一种或多种症状,使得所述受试者受到治疗性治疗。

[1515] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)的组合物经由全身施用施用于受试者的中枢神经系统。在一些实施方案中,全身施用为静脉内(IV)注射。在一些实施方案中,本文所述的AAV颗粒或包含本文所述的AAV颗粒的药物组合物通过聚焦超声波(FUS),例如与微泡的静脉内施用相结合(FUS-MB),或MRI引导的FUS与静脉内施用相结合施用。

[1516] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的

组合物经由脑室内施用施用于受试者的中枢神经系统。在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由大池内注射(ICM)施用。

[1517] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由脑室内注射和静脉内注射施用于受试者的中枢神经系统。

[1518] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由ICM注射和静脉内注射以特定剂量/受试者施用于受试者的中枢神经系统。作为非限制性实例,AAV颗粒经由ICM注射以1×10^{4}个VG/受试者的剂量施用。作为非限制性实例,AAV颗粒经由IV注射以2×10^{13}个VG/受试者的剂量施用。

[1519] 在一些实施方案中,将包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物施用于受试者的中枢神经系统。在其他实施方案中,将包含本公开的AAV颗粒的组合物施用于受试者的CNS组织(例如,受试者的壳核、海马体、丘脑或皮层)。

[1520] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由实质内注射施用于受试者的中枢神经系统。实质内注射的非限制性实例包括壳核内、皮层内、丘脑内、纹状体内、海马体内或注射到内嗅皮层中。

[1521] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由实质内注射和静脉内注射施用于受试者的中枢神经系统。

[1522] 在一些实施方案中,包含本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由脑室内注射、实质内注射和静脉内注射向受试者的中枢神经系统施用。

[1523] 在一些实施方案中,包含本公开的多个颗粒的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合物经由静脉内注射向受试者的肌肉施用。在一些实施方案中,包含本公开的多个颗粒的AAV颗粒的组合物经由肌肉内注射向受试者的肌肉施用。

[1524] 在一些实施方案中,可将本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)递送至特定类型的细胞中,所述细胞包括但不限于丘脑、海马体、内嗅、皮层、运动、感觉、兴奋性、抑制性、交感神经元或副交感神经元;神经胶质细胞,包括少突胶质细胞、星形胶质细胞和小神经胶质细胞;和/或围绕神经元的其他细胞,诸如T细胞。在一些实施方案中,可将本公开的AAV颗粒递送至肌肉细胞,例如股四头肌、横隔膜、肝和/或心脏(例如,心房或心室)的细胞中。

[1525] 在一些实施方案中,可将本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)(例如,多个颗粒)递送至中脑的细胞或区域。在一些实施方案中,可将本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)(例如,多个颗粒)递送至脑干的细胞或区域。

[1526] 在一些实施方案中,可将本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)(例如,多个颗粒)递送至壳核、海马体、丘脑和/或皮层中的神经元。

[1527] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作遗传病症(例如,常染色体显性遗传病症、常染色体隐性病症、X-连锁显性遗传病症、X-连锁隐性遗传病症或Y-连锁遗传病症)的疗法。在一些实施方案中,遗传病症为单基因病症或多基因病症。在一些实施方案中,遗传病症(例如,单基因病症)的治疗包括将本文所述的AAV颗粒(例如,包含本文所述的AAV衣壳变体的AAV颗粒)用于基因替代疗法。

[1528] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,

多个颗粒)可用作神经疾病的疗法。

[1529] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作tau蛋白病变的疗法。

[1530] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作阿尔茨海默氏病的疗法。

[1531] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作肌萎缩性侧索硬化的疗法。

[1532] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作亨廷顿氏病的疗法。

[1533] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作帕金森氏病的疗法。

[1534] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作戈谢病(GD)(例如,1型GD、2型GD或3型GD)的疗法。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作与GBA突变相关的帕金森氏病的疗法。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作路易体痴呆(DLB)的疗法。

[1535] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作脊髓性肌萎缩的疗法。

[1536] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作脑白质营养不良(例如,亚历山大病、常染色体显性脑白质营养不良伴自主神经疾病(ADLD)、卡纳万病、脑腱黄瘤病(CTX)、异染性脑白质营养不良(MLD)、佩-梅二氏病或雷夫叙姆病)的疗法。

[1537] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作慢性或神经病理性疼痛的疗法。

[1538] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作与HER2表达相关的疾病(例如,与HER2过表达相关的疾病)的疗法。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可用于治疗、预防、缓解或改善HER2阳性癌症。在一些实施方案中,HER2阳性癌症为HER2阳性实体肿瘤。另外或替代地,HER2阳性癌症可为局部晚期或转移性HER2阳性癌症。在一些情况下,HER2阳性癌症为HER2阳性乳腺癌或HER2阳性胃癌。在一些实施方案中,HER2阳性癌症选自由HER2阳性胃食道结合部癌、HER2阳性结肠直肠癌、HER2阳性肺癌(例如,HER2阳性非小细胞肺癌)、HER2阳性胰脏癌、HER2阳性结肠直肠癌、HER2阳性膀胱癌、HER2阳性唾液腺管癌、HER2阳性卵巢癌(例如,HER2阳性上皮卵巢癌)或HER2阳性子宫内膜癌组成的组。在一些情况下,HER2阳性癌症为前列腺癌。在一些实施方案中,HER2阳性癌症已转移至中枢神经系统(CNS)。在一些情况下,转移的HER2癌症已形成CNS赘瘤。

[1539] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作神经肿瘤性病症的疗法。在一些实施方案中,神经肿瘤性病症为原发性CNS起源的癌症(例如,CNS细胞和/或CNS组织的癌症)。在一些实施方案中,神经肿瘤性病症为CNS细胞、CNS区域和/或CNS组织中的转移性癌症。原发性CNS癌症的实例可能为神经胶质

瘤(其可包括胶质母细胞瘤也称为多形性胶质母细胞瘤)、星形细胞瘤、寡树突神经胶质瘤和室管膜瘤及混合性神经胶质瘤)、脑膜瘤、神经管胚细胞瘤、神经瘤和原发性CNS淋巴瘤(在脑、脊髓或脑膜中)等等。转移性癌症的实例包括起源于另一组织或器官的那些,例如乳腺癌、肺癌、淋巴瘤、白血病、黑色素瘤(皮肤癌)、结肠癌、肾癌、前列腺癌或转移至脑的其他类型。

[1540] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作肌肉病症或神经肌肉病症的疗法。

[1541] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)(例如,多个颗粒)可用作心脏病(cardiac disease/heart disease)的疗法和/或改善(例如,增强)受试者的心脏功能的方法。在一些实施方案中,心脏病为心肌病(例如,致心律失常性右心室心肌病、扩张型心肌病或肥厚性心肌病)、充血性心脏衰竭、心动过速(例如,儿茶酚胺激导性多形性室性心动过速)、缺血性心脏病和/或心肌梗塞。

[1542] 在一些实施方案中,相对于对照(例如,在接收AAV颗粒之前受试者中的基因、mRNA和/或mRNA水平),将本文所述的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)施用于受试者可提高受试者的靶基因、mRNA和/或蛋白质水平。在受试者(诸如但不限于受试者的CNS、CNS区域或CNS的特定细胞,或肌肉、肌肉区域或肌肉细胞)中,靶基因、mRNA和/或蛋白质水平可提高约30%、40%、50%、60%、70%、80%、85%、90%、95%和100%,或至少20-30%、20-40%、20-50%、20-60%、20-70%、20-80%、20-90%、20-95%、20-100%、30-40%、30-50%、30-60%、30-70%、30-80%、30-90%、30-95%、30-100%、40-50%、40-60%、40-70%、40-80%、40-90%、40-95%、40-100%、50-60%、50-70%、50-80%、50-90%、50-95%、50-100%、60-70%、60-80%、60-90%、60-95%、60-100%、70-80%、70-90%、70-95%、70-100%、80-90%、80-95%、80-100%、90-95%、90-100%或95-100%。在一些实施方案中,CNS的细胞包含星形胶质细胞、小神经胶质细胞、皮层神经元、海马体神经元、DRG和/或交感神经元、感觉神经元、少突胶质细胞、运动神经元或其组合。作为非限制性实例,AAV颗粒可使靶蛋白的基因、mRNA和/或蛋白水平增加,相对基线达倍数增加。在一些实施方案中,AAV颗粒导致靶基因、mRNA或蛋白质的水平提高5-6倍。

[1543] 在一些实施方案中,相对于对照(例如,在接收AAV颗粒之前受试者中的基因、mRNA和/或mRNA水平),将本文所述的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)(例如,包含编码siRNA分子或抗体或抗体片段的核酸的AAV颗粒)施用于受试者可减少受试者的靶基因、mRNA和/或蛋白质水平。在受试者(诸如但不限于受试者的CNS、CNS区域或CNS的特定细胞,或肌肉、肌肉区域或肌肉细胞)中,靶基因、mRNA和/或蛋白质水平可减少约30%、40%、50%、60%、70%、80%、85%、90%、95%和100%,或至少20-30%、20-40%、20-50%、20-60%、20-70%、20-80%、20-90%、20-95%、20-100%、30-40%、30-50%、30-60%、30-70%、30-80%、30-90%、30-95%、30-100%、40-50%、40-60%、40-70%、40-80%、40-90%、40-95%、40-100%、50-60%、50-70%、50-80%、50-90%、50-95%、50-100%、60-70%、60-80%、60-90%、60-95%、60-100%、70-80%、70-90%、70-95%、70-100%、80-90%、80-95%、80-100%、90-95%、90-100%或95-100%。在一些实施方案中,CNS的细胞包含星形胶质细胞、小神经胶质细胞、皮层神经元、海马体神经元、DRG和/或交感神经元、感觉神经元、少突胶质细胞、运动神经元或其组合。作为非限制性实例,AAV颗粒可

使靶蛋白的基因、mRNA和/或蛋白水平减少,相对基线达倍数减少。

[1544] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)可用于增加靶蛋白且减轻受试者的神经疾病的症状。在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)可用于减少靶蛋白且减轻受试者的神经疾病的症状。

[1545] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)可用于减少功能容量和日常生活活动的下降,如通过标准评估系统(诸如但不限于总功能容量(TFC)量表)所测量。

[1546] 在一些实施方案中,本公开的AAV颗粒(例如,包含AAV衣壳多肽,例如AAV衣壳变体的AAV颗粒)可用于改善用于测量神经疾病的症状的任何评估的效能。此类评估包括但不限于ADAS-cog(阿尔兹海默病评估量表-认知)、MMSE(简易精神状态检查)、GDS(老年抑郁量表)、FAQ(功能活动问卷)、ADL(日常生活活动)、GPCOG(全科医师认知评估)、Mini-Cog、AMTS(简略精神测试评分)、画钟测试、6-CIT(6项认知障碍测试)、TYM(测试你的记忆力)、MoCa(蒙特利尔认知评估)、ACE-R(Addenbrookes认知评估)、MIS(记忆障碍筛查)、BADLS(布里斯托尔日常生活活动量表)、巴式指数、功能独立性测量、工具性日常生活活动、IQCODE(老年人认知衰退知情人问卷)、神经精神问卷、Cohen-Mansfield激越问卷、BEHAVE-AD、EuroQol、简表36和/或MBR护理人员应变仪,或如Sheehan B(Ther Adv Neurol Disord.5(6):349-358(2012))中所述的其他测试中的任一种,其内容以引用方式整体并入本文中。

[1547] 在一些实施方案中,本发明组合物作为单独治疗剂或作为组合治疗剂施用以用于治疗神经疾病/病症或神经退行性病症、肌肉病症或神经肌肉病症和/或神经肿瘤性病症。

[1548] 编码靶蛋白的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可与一种或多种其他治疗剂组合使用。在一些实施方案中,组合物可与额外治疗或医疗程序同时、在额外治疗或医疗程序之前或之后施用。一般来说,各剂将以针对彼剂确定的剂量和/或时程施用。

[1549] 可与本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)组合使用的治疗剂可为小分子化合物,其为抗氧化剂、消炎剂、抗凋亡剂、钙调节剂、抗谷氨酸激导剂、结构蛋白抑制剂、参与肌肉功能的化合物以及参与金属离子调节的化合物。作为非限制性实例,所述组合疗法可与一种或多种神经保护剂(诸如小分子化合物、生长因子和激素)组合,这些神经保护剂已针对其对运动神经元变性的神经保护效应进行测试。

[1550] 经测试可与本文所述的AAV颗粒组合使用以治疗神经疾病的化合物包括但不限于胆碱酯酶抑制剂(多奈哌齐、卡巴拉汀、加兰他敏)、NMDA受体拮抗剂(诸如美金刚)、抗精神病药、抗抑郁药、抗惊厥药(例如,用于肌阵挛的丙戊酸钠和左乙拉西坦)、分泌酶抑制剂、淀粉样蛋白聚集抑制剂、铜或锌调节剂、BACE抑制剂、tau聚集抑制剂(诸如亚甲蓝、吩噻嗪、蒽醌、n-苯胺或若丹明)、微管稳定剂(诸如NAP、紫杉醇或太平洋紫杉醇)、激酶或磷酸酶抑制剂(诸如靶向GSK3β(锂)或PP2A的那些)、用Aβ肽或tau磷酸表位免疫、抗tau或抗淀粉样蛋白抗体、多巴胺耗竭剂(例如,四苯喹嗪用于舞蹈病)、苯二氮平(例如,用于肌阵挛、舞蹈病、肌肉紧张不足、强直和/或痉挛的可那氮平)、多巴胺的氨基酸前体(例如,用于强直的左旋多巴)、骨骼肌松弛剂(例如,用于强直和/或痉挛的巴氯芬、替扎尼定)、神经肌肉结合部释放乙酰胆碱导致肌肉麻痹的抑制剂(例如,用于磨牙症和/或肌肉紧张不足的肉毒杆菌毒素)、非典型抗精神病药(例如,用于精神病和/或应激性的奥氮平和喹硫平、用于精神病、舞蹈病

和/或应激性的利培酮、舒必利和氟哌啶醇、用于难治性精神病的氯氮平、用于具有突出阴性症状的精神病的阿立哌唑)、选择性5-羟色胺再摄取抑制剂(SSRI)(例如,用于抑郁症、焦虑症、强迫行为和/或应激性的西酞普兰、氟西汀、帕罗西汀、舍曲林、米氮平、文拉法辛)、催眠药(例如,用于改变睡眠-觉醒周期的索匹克隆和/或唑吡坦)、抗惊厥药(例如,用于躁狂或轻躁狂的丙戊酸钠和卡马西平)和情绪稳定剂(例如,用于躁狂或轻躁狂的锂)。

[1551] 神经营养因子可用于本公开的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)的组合疗法以治疗神经疾病。一般来说,神经营养因子被定义为促进神经元的存活、生长、分化、增殖和/或成熟,或刺激神经元的活性增加的物质。在一些实施方案中,本发明方法进一步包括将一种或多种营养因子递送至需要治疗的受试者中。营养因子可包括但不限于IGF-I、GDNF、BDNF、CTNF、VEGF、Colivelin、沙利罗登(Xaliproden)、促甲状腺素释放激素和ADNF及其变体。

[1552] 一方面,本文所述的AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)可与表达神经营养因子的AAV颗粒共施用,这些神经营养因子诸如AAV-IGF-I(参见,例如Vincent等人,Neuromolecular medicine,2004,6,79-85;其内容以引用方式整体并入本文中)和AAV-GDNF(参见,例如Wang等人,J Neurosci.,2002,22,6920-6928;其内容以引用方式整体并入本文中)。

[1553] 在一些实施方案中,向受试者施用AAV颗粒(例如,包含AAV衣壳变体的AAV颗粒)将调节(例如,增加或减少)靶蛋白在受试者中的表达,且靶蛋白的存在、水平、活性和/或表达的调节(例如,增加或减少)将减轻受试者的神经疾病/病症或神经退行性病症、肌肉病症或神经肌肉病症和/或神经肿瘤性病症的效应和/或症状。

[1554] 定义

[1555] 除非另有定义,否则本文所用的所有技术和科学术语均具有与本发明的相关领域的一般技术者通常所理解相同的含义。

[1556] 除非指示相反情形或从上下文另外显而易见,诸如“一(a/an)”和“所述(the)”的冠词可意谓一个(种)或超过一个(种)。除非指示相反情形或从上下文另外显而易见,如果一个、超过一个或全部组成员存在于、用于既定产品或过程中或以其他方式与既定产品或过程相关,则在组的一个或多个成员之间包括“或”的技术方案或描述被视为得到满足。本公开包括其中组中恰好有一个成员存在于、用于既定产品或过程中或以其他方式与既定产品或过程相关的实施方案。本公开包括其中超过一个或全部组成员存在于、用于既定产品或过程中或以其他方式与既定产品或过程相关的实施方案。

[1557] 还应注意,术语“包含”意图为开放式的且允许但不需要包括额外要素或步骤。当本文中使用术语“包含”时,因此还涵盖且公开术语“由.....组成”和“基本上由.....组成”。

[1558] 当给出范围时,包括终点。此外,应理解除非另外指示或自上下文及一般技术者的理解另外显而易见,呈范围表述的值可在本公开的不同实施方案中假定在规定范围内的任何特定值或子范围,除非上下文另外清楚指示,否则达到该范围的下限的单位的十分之一。

[1559] 腺相关病毒:如本文所用,术语“腺相关病毒”或“AAV”是指依赖病毒属的成员或其变体(例如,功能变体)。在一些实施方案中,AAV为野生型或天然存在的。在一些实施方案中,AAV为重组的。

[1560] AAV颗粒:如本文所用,“AAV颗粒”是指AAV衣壳(例如,AAV衣壳变体),及多核苷酸(例如,病毒基因组)。在一些实施方案中,AAV颗粒的病毒基因组包含至少一个有效载荷区域和至少一个ITR。在一些实施方案中,本公开的AAV颗粒为包含AAV变体的AAV颗粒。在一些实施方案中,AAV颗粒能够将编码有效载荷的核酸(例如,有效载荷区域)递送至细胞,典型地为哺乳动物(例如,人类)细胞。在一些实施方案中,本公开的AAV颗粒可重组产生。在一些实施方案中,AAV颗粒可源自本文所述或本领域中已知的任何血清型,包括血清型的组合(例如,“假型”AAV),或源自各种基因组(例如,单链或自身互补)。在一些实施方案中,AAV颗粒可能为复制缺陷型和/或靶向型。应理解,对本公开的AAV颗粒的提及还包括其药物组合物,即使未明确陈述。

[1561] 改善:如本文所用,术语“改善(amelioration)”或“改善(ameliorating)”是指疾患或疾病的至少一种指标的严重性减轻。例如,在神经退行性病症的情况下,改善包括神经元损失的减少。

[1562] 反义链:如本文所用,术语siRNA分子的“反义链”或“第一链”或“引导链”是指与靶向用于沉默的基因的mRNA的约10-50个核苷酸(例如,约15-30、16-25、18-23或19-22个核苷酸)的区段大体上互补的链。反义链或第一链具有与所需靶mRNA序列充分互补以指导靶标特异性沉默的序列,例如,足以通过RNAi机制或过程触发所需靶mRNA的破坏的互补性。

[1563] 大约:如本文所用,如应用于一个或多个相关值的术语“大约”或“约”是指与规定的参考值相似的值。当提及诸如量、时距及其类似参数的可测量值时,该术语意图涵盖偏离规定值±20%或在一些情况下±10%或在一些情况下±5%或在一些情况±1%或在一些情况下±0.1%的变化,因为此类变化适合执行所公开的方法。

[1564] 衣壳:如本文所用,术语“衣壳”是指病毒颗粒(例如,AAV颗粒)的外部(例如,蛋白质壳),其大体上(例如,>50%、>90%或100%)为蛋白质。在一些实施方案中,衣壳为AAV衣壳,其包含本文所述的AAV衣壳蛋白,例如VP1、VP2和/或VP3多肽。AAV衣壳蛋白可为野生型AAV衣壳蛋白或变体,例如野生型或参考衣壳蛋白的结构和/或功能变体,本文中称为“AAV衣壳变体”。在一些实施方案中,本文所述的AAV衣壳变体具有封闭(例如,囊封)病毒基因组的能力和/或能够进入细胞(例如,哺乳动物细胞)中。在一些实施方案中,与野生型AAV衣壳(例如,相应野生型衣壳)的趋向性相比,本文所述的AAV衣壳变体可具有修饰的趋向性。

[1565] 互补及大体上互补:如本文所用,术语“互补”是指多核苷酸彼此形成碱基对的能力。碱基对典型地由反平行多核苷酸链中的核苷酸单元之间的氢键形成。互补多核苷酸链可以Watson-Crick方式(例如,A至T、A至U、C至G)或以允许形成双链体的任何其他方式形成碱基对。如本领域技术人员所知晓,当使用RNA而非DNA时,尿嘧啶而非胸腺嘧啶是被视为与腺嘌呤互补的碱基。然而,除非另有说明,否则当U在本公开的上下文中表达时,暗示能够取代T。完美互补性或100%互补性是指如下情形,其中一个多核苷酸链的每个核苷酸单元可与第二个多核苷酸链的核苷酸单元形成氢键。不完美互补性是指如下情形,其中两个链的一些但非全部核苷酸单元可彼此形成氢键。例如,对于两个20聚体,如果每个链上仅两个碱基对可彼此形成氢键,则这些多核苷酸链展现10%互补性。在同一实例中,如果每个链上的18个碱基对可彼此形成氢键,则这些多核苷酸链展现90%互补性。如本文所用,术语“互补”可涵盖完全互补、部分互补或大体上互补。如本文所用,术语“大体上互补”意谓siRNA具有足以结合所需靶mRNA且触发靶mRNA的RNA沉默的序列(例如,在反义链中)。“完全互补”、“完

美互补性”或“100％互补性”是指如下情形,其中一个多核苷酸或寡核苷酸链的每个核苷酸单元可与第二个多核苷酸或寡核苷酸链的核苷酸单元进行碱基配对。

[1566] 囊封:如本文所用,术语“囊封”意谓封闭、围绕或包封。例如,衣壳蛋白(例如,AAV衣壳变体)通常囊封病毒基因组。在一些实施方案中,囊封于衣壳(例如,AAV衣壳变体)内涵盖衣壳的100％覆盖,以及小于100％覆盖,例如95％或更少。例如,衣壳中可存在间隙或不连续性,只要病毒基因组经保留于衣壳中,例如在进入细胞中之前。

[1567] 有效量:如本文所用,术语剂的“有效量”为足以产生有益或所需结果(例如,临床结果)的量且因而,“有效量”取决于其中应用该量的情况。例如,在施用治疗癌症的剂的情况下,剂的有效量是如与不施用该剂所获得的反应相比,例如足以实现癌症的如本文所定义的治疗的量。

[1568] 表达:如本文所用,核酸序列的“表达”是指由DNA序列产生RNA模板(例如,通过转录)。在一些实施方案中,表达进一步包括以下一个或多个:(1)RNA转录本的加工(例如,通过剪接、编辑、5′帽形成和/或3′端加工);(2)将RNA转译成多肽或蛋白质;及(3)多肽或蛋白质的转译后修饰。

[1569] 片段:如本文所用,“片段”是指一部分。例如,抗体片段可包含CDR,或重链可变区,或scFv等。在一些实施方案中,片段为核酸片段。

[1570] 同一性:如本文所用,“同一性”是指两个聚合物分子之间(例如两个核酸分子之间,诸如两个DNA分子或两个RNA分子之间,或两个多肽分子之间)的亚单位序列同一性。当这两个分子两者中的亚单位位置由相同单体亚单位占据时;例如,如果两个DNA分子中的每个的位置均由腺嘌呤占据,则其在那个位置处一致。两个序列之间的同一性为匹配位置的数目的直接函数;例如,如果两个序列中的一半位置(例如,十个亚单位的长度的聚合物中的五个位置)同一,则这两个序列为50％同一;如果90％的位置(例如,10个中的9个)匹配,则这两个序列为90％同一。举例来说,可通过比对两个多核苷酸序列来执行这两个序列的同一性百分比的计算,以实现最佳比较目的(例如,可在第一和第二核酸序列中的一个或两个中引入间隙以实现最佳比对且出于比较目的,可忽略非一致序列)。在某些实施方案中,出于比较目的而比对的序列的长度为参考序列的长度的至少30％、至少40％、至少50％、至少60％、至少70％、至少80％、至少90％、至少95％或100％。接着比较在对应核苷酸位置处的核苷酸。当第一序列中的位置由与第二序列中的相应位置相同的核苷酸占据时,则这些分子在那个位置处同一。两个序列之间的同一性百分比为这些序列所共享的同一位置的数目的函数,其中考虑到间隙的数目和每个间隙的长度,需要引入该长度以实现两个序列的最佳比对。可使用数学算法来完成序列的比较及两个序列之间的同一性百分比的确定。例如,可使用诸如以下文献中所述的那些的方法来确定两个核苷酸序列之间的同一性百分比:Computational Molecular Biology,Lesk,A.M.编,Oxford University Press,New York,1988;Biocomputing:Informatics and Genome Projects,Smith,D.W.编,Academic Press,New York,1993;Sequence Analysis in Molecular Biology,von Heinje,G.,Academic Press,1987;Computer Analysis of Sequence Data,第I部分,Griffin,A.M.和Griffin,H.G.编,Humana Press,New Jersey,1994;及Sequence Analysis Primer,Gribskov,M.和Devereux,J.编,M Stockton Press,New York,1991;其各自的内容以引用方式整体并入本文中。例如,可使用Meyers和Miller(CABIOS,1989,4:11-17)的算法来确定

两个核苷酸序列之间的同一性百分比,该算法已并入使用PAM120权重残基表、间隙长度罚分12及间隙罚分4的ALIGN程序(2.0版)中。或者,可使用GCG软件包中的GAP程序使用NWSgapdna.CMP矩阵来确定两个核苷酸序列之间的同一性百分比。通常用于确定序列之间的同一性百分比的方法包括但不限于Carillo,H.和Lipman,D.,SIAM J Applied Math.,48:1073(1988)中所公开的那些;以引用方式并入本文中。用于确定同一性的技术经编入公开可得的计算器程序中。确定两个序列之间的同源性的例示性计算器软件包括但不限于GCG程序包(Devereux,J.等人,Nucleic Acids Research,12(1),387(1984))、BLASTP、BLASTN和FASTA(Altschul,S.F.等人,J.Molec.Biol.,215,403(1990))。

[1571] 抑制基因表达:如本文所用,措辞“抑制基因表达”意谓导致基因的表达产物的量减少。表达产物可为由该基因转录的RNA(例如,mRNA),或由该基因转录的mRNA所转译的多肽。典型地,mRNA水平的降低导致由其转译的多肽的水平的降低。可使用用于测量mRNA或蛋白质的标准技术来确定表达水平。

[1572] 分离:如本文所用,术语“分离”是指物质或实体由天然状态改变或移出,例如由在天然状态下与其缔合的组分中的至少一些改变或移出。例如,天然存在于活动物中的核酸或肽并非“分离的”,但与其天然状态的共存物质部分或完全分离的相同核酸或肽为“分离的”。分离的核酸或蛋白质可以大体上纯化的形式存在,或可存在于非原生环境(例如宿主细胞)中。此类多核苷酸可能为载体的一部分和/或此类多核苷酸或多肽可能为组合物的一部分,且仍为分离的,因为此类载体或组合物并非自然界中发现所述载体或组合物的环境的一部分。在一些实施方案中,分离的核酸为重组的或可并入载体中。

[1573] 神经疾病:如本文所用,“神经疾病”为与中枢或外周神经系统及其成分(例如,神经元)相关的任何疾病。

[1574] 正交演化:如本文所用,术语“正交演化”是指一种方法,其中在任何数目的可来自不同物种和/或菌种的细胞和/或受试者类型的集合中施用AAV颗粒以进行如本文所述的第一轮AAV选择,且其中在任何数目的可来自不同物种和/或菌种的细胞和/或受试者类型的集合中,或在任何数目的可来自同一物种和/或菌株的细胞和/或受试者类型的集合中执行任何数目的额外(例如,后续)AAV选择轮。

[1575] 有效载荷区域:如本文所用,“有效载荷区域”为编码本公开的一种或多种“有效载荷”的任何核酸序列(例如,在病毒基因组内)。作为非限制性实例,有效载荷区域可为AAV颗粒的病毒基因组内的核酸序列,其编码有效载荷,其中有效载荷为RNAi剂或多肽。本公开的有效载荷可为但不限于肽、多肽、蛋白质、抗体、RNAi剂等。

[1576] 多肽:如本文所用,“多肽”意谓最常通过肽键连接在一起的氨基酸残基(天然或非天然)的聚合物。如本文所用,该术语是指任何大小、结构或功能的蛋白质、多肽和肽。如果所述多肽为肽,则其将为至少约2、3、4个或至少5个氨基酸残基长。因此,多肽包括基因产物、天然存在的多肽、合成多肽、前述各物的同源物、异种同源物、同种同源物、片段及其他等效物、变体及类似物。多肽可为单一分子或可为多分子复合物,诸如二聚体、三聚体或四聚体。其还可包含单链或多链多肽且可缔合或连接。术语多肽也可应用于氨基酸聚合物,其中一个或多个氨基酸残基为相应天然存在的氨基酸的人工化学类似物。

[1577] 多肽变体:术语“多肽变体”是指其氨基酸序列不同于原生或参考序列的分子。如与原生或参考序列相比,这些氨基酸序列变体可具有氨基酸序列内的某些位置处的取代、

缺失和/或插入。在一些实施方案中,变体包含与原生或参考序列具有至少约50%、至少约80%或至少约90%同一(同源)的序列。

[1578] 肽:如本文所用,"肽"为小于或等于50个氨基酸长,例如约5、10、15、20、25、30、35、40、45或50个氨基酸长。

[1579] 药学上可接受:词组"药学上可接受"在本文中用于指在合理医学判断的范围内适合与人和动物的组织接触使用而无过度毒性、刺激、过敏反应或其他问题或并发症且与合理益处/风险比相称的那些化合物、材料、组合物和/或剂型。

[1580] 预防:如本文所用,术语"预防(preventing)"或"预防(prevention)"是指部分或完全地延迟感染、疾病、病症和/或疾患的发作;部分或完全地延迟特定感染、疾病、病症和/或疾患的一种或多种症状、特征或临床表现的发作;部分或完全地延迟特定感染、疾病、病症和/或疾患的一种或多种症状、特征或表现的发作;部分或完全地延迟自感染、特定疾病、病症和/或疾患进展;和/或降低发展与感染、疾病、病症和/或疾患相关的病理的风险。

[1581] 区域:如本文所用,术语"区域"是指区带或一般区域。在一些实施方案中,当提及蛋白质或蛋白质模块时,区域可包含沿着蛋白质或蛋白质模块的线性氨基酸序列或可包含三维区域、表位和/或表位簇。在一些实施方案中,区域包含末端区域。如本文所用,术语"末端区域"是指位于给定剂的末端(ends/termini)处的区域。当提及蛋白质时,末端区域可包含N端和/或C端。

[1582] 在一些实施方案中,当提及多核苷酸时,区域可包含沿着所述多核苷酸的线性核酸序列或可包含三维区域、二级结构或三级结构。在一些实施方案中,区域包含末端区域。如本文所用,术语"末端区域"是指位于给定剂的末端(ends/termini)的区域。当提及多核苷酸时,末端区域可包含5'和/或3'端。

[1583] RNA或RNA分子:如本文所用,术语"RNA"或"RNA分子"或"核糖核酸分子"是指核糖核苷酸的聚合物;术语"DNA"或"DNA分子"或"脱氧核糖核酸分子"是指脱氧核糖核苷酸的聚合物。DNA和RNA可天然合成,例如分别通过DNA复制和DNA转录;或以化学方式合成。DNA和RNA可为单链的(即分别为ssRNA或ssDNA)或多链的(例如双链,即分别为dsRNA和dsDNA)。如本文所用,术语"mRNA"或"信使RNA"是指编码一个或多个多肽链的氨基酸序列的单链RNA。

[1584] RNA干扰或RNAi:如本文所用,术语"RNA干扰"或"RNAi"是指由RNA分子介导的序列特异性调节机制,其导致相应蛋白质编码基因的表达的抑制或干扰或"沉默"。已在包括植物、动物和真菌在内的多种类型的生物体中观察到RNAi。RNAi天然地出现于细胞中以移除外来RNA(例如,病毒RNA)。天然RNAi经由自游离dsRNA裂解的片段进行,这些片段将降解机制引导至其他类似RNA序列。RNAi由RNA诱导的沉默复合物(RISC)控制,且由细胞质中的短/小dsRNA分子起始,其中dsRNA分子与催化性RISC组分argonaute相互作用。dsRNA分子可外源性地引入细胞中。外源dsRNA通过活化核糖核酸酶蛋白Dicer起始RNAi,Dicer结合且裂解dsRNA以产生21-25个碱基对的双链片段,其中每端具有一些未配对的悬垂碱基。这些短双链片段被称为小干扰RNA(siRNA)。

[1585] RNAi剂:如本文所用,术语"RNAi剂"是指RNA分子或其衍生物,其可诱导靶基因和/或其蛋白质产物的表达的抑制、干扰或"沉默"。RNAi剂可敲除(实际上消除或消除)表达,或敲低(减少或降低)表达。RNAi剂可为但不限于dsRNA、siRNA、shRNA、前体miRNA、初级miRNA、miRNA、stRNA、lncRNA、piRNA或snoRNA。

[1586] miR结合位点:如本文所用,“miR结合位点”包含核酸序列(RNA或DNA,例如差别在于RNA的“U”或DNA的“T”),该核酸序列能够全部或部分地与微小RNA(miR)结合,或全部或部分地与微小RNA(miR)结合,例如经由完全或部分杂交。典型地,此类结合发生于miR与miR结合位点之间,呈反向补体取向。在一些实施方案中,miR结合位点由编码miR结合位点的AAV病毒基因组转录。

[1587] 在一些实施方案中,miR结合位点可连续地编码或转录。此类“miR结合位点系列”或“miR BS”可包括两个或两个以上具有相同或不同核酸序列的miR结合位点。

[1588] 间隔子:如本文所用,“间隔子”通常为任何所选择的核酸序列,例如1、2、3、4、5、6、7、8、9或10个核苷酸的长度,该序列位于两个或两个以上连续的miR结合位点序列之间。间隔子也可为超过10个核苷酸的长度,例如20、30、40或50个或超过50个核苷酸。

[1589] 样品:如本文所用,术语“样品”或“生物样品”是指其组织、细胞、核酸或组分部分(例如体液,包括但不限于血液、血清、粘液、淋巴液、滑液、脑脊液、唾液、羊水、羊水脐带血、尿液、阴道液和精液)的子集。

[1590] 自身互补病毒颗粒:如本文所用,“自身互补病毒颗粒”是包含至少两种组分(蛋白质衣壳和封入所述衣壳内的自身互补病毒基因组)的颗粒。

[1591] 有义链:如本文所用,术语siRNA分子的“有义链”或“第二链”或“过客链”是指与反义链或第一链互补的链。siRNA分子的反义链与有义链杂交形成双链体结构。如本文所用,“siRNA双链体”包括与靶向用于沉默的基因的mRNA的约10-50个核苷酸的区段具有足够互补性的siRNA链,及具有足够互补性以与另一siRNA链形成双链体的siRNA链。

[1592] 短干扰RNA或siRNA:如本文所用,术语“短干扰RNA”、“小干扰RNA”或“siRNA”是指包含约5-60个之间核苷酸(或核苷酸类似物)的RNA分子(或RNA类似物),其能够指导或介导RNAi。优选地,siRNA分子包含约15-30个之间核苷酸或核苷酸类似物,诸如约16-25个之间核苷酸(或核苷酸类似物)、约18-23个之间核苷酸(或核苷酸类似物)、约19-22个之间核苷酸(或核苷酸类似物)(例如,19、20、21或22个核苷酸或核苷酸类似物)、约19-25个之间核苷酸(或核苷酸类似物)以及约19-24个之间核苷酸(或核苷酸类似物)。术语“短”siRNA是指包含5-23个核苷酸、优选21个核苷酸(或核苷酸类似物)(例如,19、20、21或22个核苷酸)的siRNA。术语“长”siRNA是指包含24-60个核苷酸、优选约24-25个核苷酸)(例如,23、24、25或26个核苷酸)的siRNA。在一些情况下,短siRNA可包括少于19个核苷酸,例如16、17或18个核苷酸,或少至5个核苷酸,其限制条件在于较短siRNA保留介导RNAi的能力。同样,在一些情况下,长siRNA可包括超过26个核苷酸,例如27、28、29、30、35、40、45、50、55或甚至60个核苷酸,其限制条件在于较长siRNA保留在未进一步加工(例如,酶加工)成短siRNA的情况下介导RNAi或转译抑制的能力。siRNA可为单链RNA分子(ss-siRNA)或包含有义链和反义链的双链RNA分子(ds-siRNA),这些链杂交形成称为siRNA双链体的双链体结构。

[1593] 受试者:如本文所用,术语“受试者”或“患者”是指可施用根据本公开的组合物的任何生物体,例如以达成实验、诊断、预防和/或治疗目的。典型受试者包括动物(例如哺乳动物,诸如小鼠、大鼠、兔、非人灵长类动物和人)和/或植物。

[1594] 靶细胞:如本文所用,“靶细胞”或“靶组织”是指任何一种或多种相关细胞。这些细胞可发现于体外、体内、原位或生物体的组织或器官中。生物体可为动物,优选哺乳动物,更优选人且最优选患者。

[1595] 治疗剂:术语“治疗剂”是指当施用于受试者时具有治疗、诊断和/或预防效应和/或引发所需生物学和/或药理学效应的任何剂。

[1596] 治疗有效量:如本文所用,术语“治疗有效量”意谓有待递送的剂(例如,核酸、药物、治疗剂、诊断剂、预防剂等)的量,当施用于罹患或易患感染、疾病、病症和/或疾患的受试者时,该量足以对所述感染、疾病、病症和/或疾患进行治疗、改善其症状、进行诊断、预防和/或延迟其发作。在一些实施方案中,治疗有效量以单一剂量提供。

[1597] 治疗:如本文所用,术语“治疗”是指部分地或完全地减轻、改善、改良、缓解特定感染、疾病、病症和/或疾患的一种或多种症状或特征、延迟其发作、抑制其进展、减轻其严重性和/或降低其发生率。例如,“治疗”癌症可指抑制肿瘤的存活、生长和/或扩散。可将治疗施用于未展现疾病、病症和/或疾患的迹象的受试者和/或仅展现疾病、病症和/或疾患的早期迹象的受试者,以达成降低发展与所述疾病、病症和/或疾患相关的病理的风险的目的。

[1598] 保守氨基酸取代:如本文所用,“保守氨基酸取代”是其中氨基酸残基经具有相似侧链的氨基酸残基置换的取代。具有相似侧链的氨基酸残基的家族已在本领域中定义。这些家族包括具有以下侧链的氨基酸:碱性侧链(例如,赖氨酸、精氨酸、组氨酸)、酸性侧链(例如,天冬氨酸、谷氨酸)、不带电极性侧链(例如,甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸)、非极性侧链(例如,丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸)、β分支链侧链(例如,苏氨酸、缬氨酸、异亮氨酸)和芳族侧链(例如,酪氨酸、苯丙氨酸、色氨酸、组氨酸)。

[1599] 变体:如本文所用,术语“变体”是指具有与参考序列大体上同一,例如具有至少70%、75%、80%、85%、90%、95%或99%序列同一性的氨基酸或核苷酸序列的多肽或多核苷酸。在一些实施方案中,变体为功能变体。

[1600] 功能变体:如本文所用,术语“功能变体”是指具有参考序列的至少一种活性的多肽变体或多核苷酸变体。

[1601] 载体:如本文所用,术语“载体”是指转运、转导或以其他方式充当异源分子的载剂的任何分子或部分。在一些实施方案中,载体可为质体。本公开的载体可重组产生。异源分子可为多核苷酸和/或多肽。

[1602] 病毒基因组:如本文所用,术语“病毒基因组”是指囊封于AAV颗粒中的核酸序列。病毒基因组包含具有至少一个编码有效载荷的有效载荷区域和至少一个ITR的核酸序列。

[1603] 等效物和范围

[1604] 本文所引用的每项专利、专利申请和公布的公开内容均由此以引用方式整体并入本文中。本领域技术人员将认识到或能够仅使用常规实验来确定根据本文所述的本公开的特定实施方案的许多等效物。本公开的范围不意图局限于上述说明书,而是如随附权利要求书中所阐述。

[1605] 另外,应理解属于先前技术的本公开的任何特定实施方案均可自权利要求中的任一项或多项排除。因为认为此类实施方案是一般技术者已知的,故即使本文未明确地阐述该排除,也可排除这些实施方案。本公开的组合物的任何特定实施方案(例如,任何抗生素、治疗性或活性成分;任何制造方法;任何使用方法;等)可出于任何原因从任一个或多个技术方案中排除,无论是否与先前技术的存在相关。

[1606] 应理解,已使用的措辞为描述而非限制措辞,且可在随附权利要求书的范围内作

出变化,而不偏离本公开在其更宽方面中的真实范围和精神。

[1607] 虽然已关于数个所述实施方案以某一长度且以某一特殊性描述了本公开,但不意图应将其局限于任何此类详情或实施方案或任何特定实施方案,而应参考随附权利要求书对其进行解释以鉴于先前技术对此类技术方案提供最广泛的可能解释且因此有效地涵盖本公开的预期范围。

[1608] 通过以下非限制性实施例进一步说明本公开。

[1609] 实施例

[1610] 实施例1.NHP和大鼠中TRACER AAV文库的高通量筛选

[1611] 使用如WO 2020/072683、WO 2021/202651和WO 2021/230987(其内容以引用方式整体并入本文中)中所述的基于TRACER的方法来生成本文所述的AAV衣壳变体。使正交演化方法与通过NGS进行的高通量筛选组合。简单来说,使用突变诱发方法生成AAV衣壳变体的文库,其中相对于根据SEQ ID NO:138编号的参考序列,将7至8个氨基酸的长度的序列插入AAV5的环VIII中的不同位置中,包括残基570-584之间。使初始文库经由非人灵长类动物(NHP)、具体地食蟹猕猴(cynomolgus macaque/Macaca fascicularis)、大鼠或人类脑微血管内皮细胞(hBMVEC)传代三次。在每个系统中的第三代之后,将来自NHP的572种变体、来自大鼠的80种变体及来自hBMVEC的99种变体汇集至具有747种总变体的第3代合成文库中。接着使此文库在NHP和大鼠中传代。在此传代之后(例如,注射到两个NHP和大鼠中后一个月),从三个脑区提取RNA。在RNA回收和RT-PCR扩增之后,执行系统性NGS富集分析以计算相对于AAV5野生型对照在NHP(表9,左栏)和大鼠(表9,右栏)中的倍数富集,且在两种动物中均鉴定出包含于所述变体内的肽。高于1的倍数富集值指示相对于AAV5的表达增加。在不同宿主中生成且传代的AAV5变体的所有文库均处于突触蛋白启动子的控制下。

[1612] 如表9所示,在NHP的脑中鉴定出大约288种变体,其平均倍数变化大于野生型AAV5。在288种NHP变体中,27种证实与野生型相比大于5的倍数变化,其中1种变体证实大于60的倍数变化。例如,包含YPAEVVQK(SEQ ID NO:943)的变体在NHP的脑中证实64.9倍富集。

[1613] 如表9所示,在大鼠的脑中鉴定出大约98种变体,其平均倍数变化大于野生型AAV5。在98种变体中,33种证实与野生型相比大于2的倍数变化,其中1种变体证实大于40的倍数变化。例如,包含YPAEVVQK(SEQ ID NO:943)的变体在大鼠的脑中证实41.1倍富集。

[1614] 包含YPAEVVQK(SEQ ID NO:943)的变体相对于野生型AAV5在NHP的脑中证实高倍数富集(64.9倍富集),在大鼠的脑中也证实高倍数变化(41.1倍富集)。这指示包含SEQ ID NO:943的此AAV衣壳变体能够跨物种,如NHP和大鼠脑中的表达和趋向性增加所证明。

[1615] 表9.NHP和大鼠中AAV衣壳变体的NGS倍数富集

[1616]

肽序列	SEQ ID NO:	NHP 的脑中相对于 AAV5 的倍数富集	肽序列	SEQ ID NO:	大鼠的脑中相对于 AAV5 的倍数富集
YPAEVVQK	943	64.910	YPAEVVQK	943	41.054
RNQKVQGG	202	23.240	NHVVVISG	432	25.930
PDNRVKPV	203	17.843	TYAKDTSM	659	16.868
ARVEVSAG	204	12.970	TVLAQSTP	547	7.098
MLKRVITI	205	10.323	TFQTANKV	235	6.518
ESTHVELI	206	10.126	RAQTVSVK	638	5.209
GRVHVNMA	207	9.652	TMLNSAQQ	666	4.636
TSQSSTKS	208	9.586	RLAKVTQP	223	4.558
AKQVVKPP	209	8.506	TIPNKVTA	779	3.480
STMDVNKR	210	8.284	TPEATNHK	271	3.471
KKSQVANA	211	8.264	TTLSAKDR	383	3.334
TNAVSTKQ	212	8.071	PSPNVTKN	295	3.285
TLATHENM	213	7.093	TRPNPSPK	850	3.255

[1617]

VKGKVSAG	214	6.637	TANSVKSQ	319	3.003
IRQKVPAS	215	6.467	SDNLVKHD	664	2.973
GRSEVLKG	216	6.139	LRGQVKSN	248	2.932
TPIKTAVR	217	6.088	TLKHSDLS	264	2.904
CKLQVDCQ	218	5.842	EVKMVSNV	380	2.787
TKLKPAGP	219	5.830	RNQKVQGG	202	2.769
NTMHVELR	220	5.721	RLVTVKSP	332	2.690
TKTKLSSP	221	5.686	IIAEVGKQ	528	2.519
DNMKVLGG	222	5.585	TKQMSDHG	336	2.455
RLAKVTQP	223	5.454	TIGGMKGK	390	2.435
GSNMVASK	224	5.325	SHSDVSVR	776	2.383
TMAPTKNP	225	5.241	TKTLQTSP	631	2.364
TAQHTTLS	226	5.040	VMSNVTQR	483	2.216
TQPSGKTQ	227	5.024	TKLSAVTK	755	2.213
MNSKVSSV	228	4.985	TLASHENI	329	2.198
TEATVKSN	229	4.971	TKLKPAGP	219	2.161
TSEIQGAK	230	4.969	TTAKPTVE	516	2.133
LSSNVTKA	231	4.952	TSEPPSTE	930	2.078
TPLCGMAL	232	4.830	ALNSVRRE	781	2.053
NNTAVSNK	233	4.789	LHNKVNDV	680	2.014
NKNRVRDL	234	4.770	LSSNVTKA	231	1.953
TFQTANKV	235	4.669	TSNAVRQQ	765	1.948
TTKQPKAL	236	4.662	QGANVVKQ	320	1.901
GSAEVKMR	237	4.650	KKPTVATI	300	1.892
SLLQVKL	238	4.515	KMAGVARE	615	1.872

[1618]

E					
THAVSTKP	239	4.431	TMASQEHM	661	1.857
DINSVRRG	240	4.425	GGQGVNQS	821	1.812
QSKTVPAK	241	4.181	RPQKVSSA	386	1.741
TGKPDKSY	242	4.163	NLNIVRAP	869	1.728
TKIPSTSR	243	4.055	GMNSVRNA	854	1.726
TPAATNPK	244	4.036	LSSKVTDS	1596	1.622
TITKPPAP	245	4.007	NPTVVKSP	303	1.576
GRLIVTDG	246	4.003	NSPNVSRQ	589	1.563
TPKKSTSV	247	3.994	MLNEVSRR	872	1.562
LRGQVKSN	248	3.979	ITSSVARS	891	1.561
GLPDVIKQ	249	3.920	TKIPSTSR	243	1.533
TREQTSTA	250	3.892	TKVSGTPD	267	1.516
TRNGQPVN	251	3.847	ALDAVSNK	283	1.514
TRASLSIK	252	3.807	SRENVSRQ	478	1.464
SFSEVKNM	253	3.748	EAMMVDRK	608	1.401
TQSRSGGP	254	3.666	THTRIIAT	784	1.397
TSTPAVSQ	255	3.650	LSTKVSDQ	721	1.396
TASVPATS	256	3.626	TSSITPNP	268	1.387
TLASHENK	257	3.600	TATPSQSK	335	1.373
TQSSDHKV	258	3.590	RSKSVNPM	327	1.367
TLASQEHM	259	3.590	RYMTVKTQ	2083	1.363
TKGLPTQM	260	3.581	TTMSATSP	324	1.354
SSVVVDKM	261	3.542	TSTPAVSQ	255	1.348
TLDSHEHK	262	3.516	PLKNVPPG	577	1.329
TNQKVT	263	3.511	TPIKTAVR	217	1.318

[1619]

GS					
TLKHSDLS	264	3.463	TNANQVTR	691	1.296
TTSTVSSP	265	3.431	TSSSPSNP	354	1.280
TKASHEHM	266	3.359	TNGATLGS	705	1.261
TKVSGTPD	267	3.330	KLLKVGSS	316	1.249
TSSITPNP	268	3.328	PSNNVKAL	469	1.222
AAETVKPN	269	3.279	TPLCGMAL	232	1.213
TNTQSKLS	270	3.271	YANSVKAM	477	1.207
TPEATNHK	271	3.229	KRLNVESR	448	1.195
TGGRNPGG	272	3.227	IKSGVQQK	348	1.193
TGGRVQGP	273	3.224	TTDLTTVE	431	1.173
KMSGVVGK	274	3.194	KPSKVGGV	423	1.170
KGGVVKTV	275	3.190	IKPPVPKA	883	1.161
TNGASSKL	276	3.143	PFQAVLGS	604	1.160
SSGDVKLW	277	3.108	SSPNVSRN	284	1.135
TSGTPNNF	278	3.012	NKLGVVTK	439	1.130
KRTTVSSP	279	3.010	KKSQVANA	211	1.116
TLASNENM	280	2.997	KEESVDTS	886	1.110
TKQKDLPR	281	2.987	TPVSVNTH	967	1.098
ATNVVSSK	282	2.984	TRETPSTE	507	1.095
ALDAVSNK	283	2.971	LGSNVSAR	414	1.093
SSPNVSRN	284	2.955	KAVSVDTS	579	1.088
THAASNKS	285	2.933	TPMSTSDL	688	1.069
TSSSTKSI	286	2.925	PVNVVRAS	797	1.059
EGKLVGTQ	287	2.923	VSGNVSRS	677	1.047
PNKLVGSV	288	2.903	AIEKVSVN	498	1.046

[1620]

THASLPKP	289	2.868	TLGHSAKQ	511	1.046
TPGNSARS	290	2.856	QQDSVPGQ	515	1.033
TAPKVGNM	291	2.846	DNMKVLGG	222	1.030
TRALVKPP	292	2.846	TPGNSARS	290	1.030
GLEDVKGR	293	2.823	THAVTTKQ	406	1.030
QSEKVHMG	294	2.812	NCTVVQCR	326	1.024
PSPNVTKN	295	2.783	TSNRTTLQ	561	1.013
THAQSTKP	296	2.736	TITKPPAP	245	1.010
RPKTVSTF	297	2.722	TPTPPSAQ	436	1.007
EPSEVHKM	298	2.720	TLRTNESV	801	1.000
TNQKPTLQ	299	2.665	TTAPATGT	489	1.000
KKPTVATI	300	2.661	TTRAQTTK	858	0.991
TATVQSNP	301	2.652	KSRSVSGV	896	0.977
YQSDVPNR	302	2.646	TAGTTVTP	463	0.976
NPTVVKSP	303	2.624	WSNSVSTR	546	0.961
TVGSESRP	304	2.595	TGKLPSGQ	391	0.960
TSKSAAVR	305	2.591	NSNRVGGN	313	0.957
TLANHEHM	306	2.582	TVSSPSGG	709	0.955
TLASHENM	307	2.582	DVASVSSK	602	0.955
RKPSVQTV	308	2.574	TQPSGKTQ	227	0.944
TGRDVPKL	309	2.507	IRQKVPAS	215	0.939
SQSSVKTP	310	2.492	DANSVKVV	522	0.938
ATSVVSQA	311	2.477	QAHRVPVE	2082	0.932
TGTKVESY	312	2.474	KRTTVSSP	279	0.922
NSNRVG	313	2.470	TSLQVSKI	559	0.921

[1621]

GN					
TMTRVGDG	314	2.455	AGNSVRVS	537	0.920
NNGNVSAK	315	2.413	AKQVVKPP	209	0.919
KLLKVGSS	316	2.402	LAKTVNQT	442	0.912
SQPKVSTA	317	2.402	TPLDKRIR	895	0.912
MNDKVLTN	318	2.395	RQNKVTSG	863	0.897
TANSVKSQ	319	2.373	ARVEVSAG	204	0.894
QGANVVKQ	320	2.367	THAASNKS	285	0.888
TSSVSKQP	321	2.363	KMSGVVGK	274	0.873
SQGNVAKA	322	2.354	ISNMVRSS	848	0.872
IYVAVPAA	323	2.324	TTSTVSSP	265	0.863
TTMSATSP	324	2.321	TAVKSQMP	339	0.844
TFTATSNP	325	2.315	MNSKVSSV	228	0.841
NCTVVQCR	326	2.289	TSSSTKSI	286	0.840
RSKSVNPM	327	2.267	TAKILSTE	2084	0.836
GQVDVPKL	328	2.266	TDRNTPKL	351	0.834
TLASHENI	329	2.266	MVNVVHQS	813	0.829
THNSVRDR	330	2.233	TEATVKSN	229	0.825
GQSTVGQM	331	2.231	TSQSSTKS	208	0.822
RLVTVKSP	332	2.229	LVNAVKSF	811	0.821
THEVSTKQ	333	2.218	KIKVVNTP	342	0.817
TDKKQTTS	334	2.208	VLSEVKLQ	640	0.810
TATPSQSK	335	2.206	VKGKVSAG	214	0.809
TKQMSDHG	336	2.109	ATLSVSTN	618	0.806
TLASNEHK	337	2.108	VNSKVISD	549	0.793

[1622]

NMSPVPKL	338	2.106	NAVGVAHG	754	0.791
TAVKSQMP	339	2.100	TSASKVSH	357	0.788
SNGAVGKH	340	2.092	AKDAVSGL	617	0.788
SNSKVNTG	341	2.081	TSGTVGTK	447	0.787
KIKVVNTP	342	2.068	TGLSVSTN	636	0.777
LLETVATK	343	2.060	MSNKVLRD	532	0.777
DSAKVMPQ	344	2.045	YMNSVKSF	876	0.777
TTKEVLSG	345	2.042	TLASNENM	280	0.776
ISSNVARG	346	2.037	ATNVVSSK	282	0.769
TGSTQSDL	347	2.027	TNQSGPGS	922	0.766
IKSGVQQK	348	2.023	QLNKVVIR	479	0.758
TATATQSK	349	2.017	LSNSVTPR	509	0.754
DGSQVHNR	350	1.996	TTKEVLSG	345	0.748
TDRNTPKL	351	1.984	TQSSDHKV	258	0.741
TTSGSPQP	352	1.966	YSNNVVKS	672	0.740
THEVSTKP	353	1.958	EGLKVLVG	497	0.739
TSSSPSNP	354	1.933	TTTSVNRP	746	0.739
LVAPVKQV	355	1.924	TDNRVRSY	880	0.732
LELLVPKL	356	1.922	TSSPGMKL	753	0.731
TSASKVSH	357	1.903	TTKQPKAL	236	0.726
LNAGVSSP	358	1.901	LGNRVSAL	722	0.722
TVTNGGKG	359	1.886	KLRQVNAI	481	0.715
TLGKVSLE	360	1.879	TMLSSAEV	793	0.713
NQEAVVKP	361	1.874	AVVGVGQA	936	0.711
AAMTVQ	362	1.851	TTSKNRPD	900	0.709

[1623]

KV					
KPQDVSSR	363	1.843	TKLNVTAS	502	0.705
YESKVQHT	364	1.837	TATSSHQQ	627	0.703
NMRGVEVK	365	1.835	TNMASSGK	373	0.699
TREKPSTA	366	1.816	PSSDVVSQ	374	0.696
VERSVTSN	367	1.805	IGPSVGTD	606	0.693
VGDLVHTG	368	1.803	TLPTPRQH	620	0.692
KNDHVSLS	369	1.796	STNTVRVN	487	0.688
SNEKVSMS	370	1.793	VGNRVTGN	758	0.688
TRRAQESP	371	1.782	NKNRVRDL	234	0.687
TTVKSDQS	372	1.773	TPEETNPK	711	0.686
TNMASSGK	373	1.759	YVSNVAKA	499	0.680
PSSDVVSQ	374	1.757	RSNKVRET	467	0.668
TLSSHENM	375	1.755	KSLDVSIS	518	0.666
TNPSMQSP	376	1.741	TPTPKNSH	614	0.660
GVSLVKTG	377	1.703	TSETPSTA	486	0.648
TPRNGAQP	378	1.701	TRALVKPP	292	0.647
TVSDKSMS	379	1.689	SNSKVNTG	341	0.643
EVKMVSNV	380	1.678	TLTQTVRD	918	0.641
VSAGVSKN	381	1.676	NNGNVSAK	315	0.637
LSKGVGHE	382	1.673	SRENVSRS	445	0.636
TTLSAKDR	383	1.668	ISNSVRQS	745	0.630
TMASHEHK	384	1.639	TLATQEHM	433	0.629
TKALTSGP	385	1.638	TDKQASGN	495	0.629
RPQKVSSA	386	1.636	THNSVRDR	330	0.628

[1624]

TPQGVSSP	387	1.601	RSTGVHSD	550	0.625
TKGTTSSH	388	1.590	TTALHTEE	849	0.624
TMTSVVSP	389	1.576	TMASHEHK	384	0.622
TIGGMKGK	390	1.564	QRNLVKSA	609	0.622
TGKLPSGQ	391	1.555	QDSTVSKQ	519	0.618
AQDAVSAR	392	1.551	TSPNRNQS	830	0.617
TKVPSVSP	393	1.551	TGGRNPGG	272	0.616
DKGTVERR	394	1.538	KAATVQEP	527	0.612
SKLDVNQR	395	1.538	TKALTSGP	385	0.611
VSGKVVHQ	396	1.528	TSPLANKS	496	0.609
ASNKVGTL	397	1.525	SANKVKPD	541	0.608
TKPPGQQL	398	1.523	NNTAVSNK	233	0.607
TLASQEHK	399	1.516	HTNSVRTP	599	0.599
SIGKVGVG	400	1.511	GSNDVRNT	628	0.598
TRGASNSP	401	1.510	PSTGVVNP	822	0.596
TNAVSTKP	402	1.492	RKPSVQTV	308	0.594
TVSSSPGT	403	1.481	SQRGVSNP	690	0.591
KAETVDTS	404	1.461	YFTGVKSE	979	0.581
TLAYHEHK	405	1.459	PLVGVGSP	716	0.581
THAVTTKQ	406	1.456	TMRTAQSK	2086	0.580
TLATHEHK	407	1.455	HSNMVRNA	844	0.578
TSLNDARR	408	1.445	LNAGVSSP	358	0.577
TNVQSSGK	409	1.433	KPQDVSSR	363	0.575
APKPVMSG	410	1.423	SGNMVRAA	795	0.573
TSVSNPS	411	1.408	TVQNSAST	592	0.572

[1625]

K					
TTSQTSAP	412	1.390	VLGSVSSN	416	0.569
TRQPSVTR	413	1.388	TPKKSTSV	247	0.568
LGSNVSAR	414	1.379	TGSTQSDL	347	0.567
TNPGVSNS	415	1.378	SANMVKHD	675	0.566
VLGSVSSN	416	1.378	GHRVVNNE	712	0.566
TLTSTGGN	417	1.366	SSSYVTVD	491	0.565
MIAKVRTE	418	1.359	GVSLVKTG	377	0.564
NGKQVSVH	419	1.349	GLPDVIKQ	249	0.564
TLLKSDRS	420	1.348	TTETASRG	975	0.563
TAGQVSSS	421	1.343	SQVVVSEK	452	0.561
TTMLAGGP	422	1.320	LRANVSSS	626	0.560
KPSKVGGV	423	1.313	LSSNVSRA	462	0.559
LSNNVRIK	424	1.309	GRLIVTDG	246	0.558
THAHVKLE	425	1.308	TTLSVSTP	905	0.551
VTKHVDLL	426	1.306	THSGVVIR	2085	0.547
SRAEVAQR	427	1.301	SITGVAKM	752	0.547
SCTNVASC	428	1.292	GGNLVRSV	730	0.546
TNSRKENI	429	1.288	SQGNVAKA	322	0.542
QQVSVPGP	430	1.278	NTDAVNQA	2081	0.541
TTDLTTVE	431	1.271	SKLDVNQR	395	0.536
NHVVVISG	432	1.254	THTTIKSP	446	0.536
TLATQEHM	433	1.246	GGGKVTDT	583	0.535
TPRSKPSP	434	1.238	TVNTTRMA	867	0.534
TTRKPQTT	435	1.229	TTSGSPQP	352	0.534

[1626]

TPTPPSAQ	436	1.225	HLSNVSKV	504	0.529
TQKPHPQG	437	1.213	TTPANSRV	538	0.526
TQLSMKAN	438	1.203	TLATHEHK	407	0.522
NKLGVVTK	439	1.202	SSKEVSRV	749	0.517
NQTDVPPR	440	1.200	HYSKVADP	695	0.515
THAETTKP	441	1.186	TFTATSNP	325	0.515
LAKTVNQT	442	1.185	TSVDKARV	492	0.514
TRPGVSDR	443	1.182	STNAVRSS	842	0.512
LMNRVTQY	444	1.181	KQRQVSVQ	747	0.511
SRENVSRS	445	1.177	RKTGVMSS	748	0.504
THTTIKSP	446	1.171	NHTTVGER	521	0.501
TSGTVGTK	447	1.165	LPSQVVKT	720	0.499
KRLNVESR	448	1.164	SGTTVDSM	454	0.499
TMAPPGKI	449	1.162	TRQPSVTR	413	0.497
TVHKASMQ	450	1.153	ANLKVIVK	723	0.492
TATASQTK	451	1.153	SGTEVASL	799	0.491
SQVVVSEK	452	1.149	TNGASSKL	276	0.490
GTATVTTR	453	1.147	DINSVRRG	240	0.487
SGTTVDSM	454	1.144	TKGTTSSH	388	0.481
ASTVVLEK	455	1.143	TERASSSL	694	0.481
TASATQTK	456	1.141	ADTRVSNP	544	0.480
SVLAVNKS	457	1.138	EPSEVHKM	298	0.477
SANLVKND	458	1.134	PDNRVKPV	203	0.477
RDAQVPKL	459	1.129	RGKAVANN	501	0.477
PQRAVQ	460	1.124	TRGASNSP	401	0.476

[1627]

EV					
TGKSPDQN	461	1.123	TAVVVNSV	578	0.475
LSSNVSRA	462	1.122	SLKSVSLD	630	0.474
TAGTTVTP	463	1.117	TMNPSRNP	707	0.471
TGSSPTGS	464	1.115	KSRNVPTL	653	0.467
KRNEVISK	465	1.112	TRASLSIK	252	0.467
TLASHEHK	466	1.105	SGSKVPTL	505	0.464
RSNKVRET	467	1.101	STLDVARR	710	0.462
TSEKIQVM	468	1.100	THAESTKQ	762	0.462
PSNNVKAL	469	1.098	VERSVTSN	367	0.457
TLATLSTV	470	1.098	KFAAVPGI	1595	0.453
THLTPKGN	471	1.097	TLATHENM	213	0.450
TPASVKGT	472	1.091	TSTRSSGN	731	0.449
TPGTPSAG	473	1.091	TREKPSTA	366	0.447
RKNTVVNS	474	1.090	ILSNVSSR	670	0.447
PQKSVSNL	475	1.088	TLTKTTMG	555	0.447
TRLSSSGS	476	1.083	TTRKPQTT	435	0.446
YANSVKAM	477	1.074	FLNAVRGN	717	0.443
SRENVSRQ	478	1.067	SVLAVNKS	457	0.442
QLNKVVIR	479	1.046	AAMTVQKV	362	0.440
TVKEKGMQ	480	1.043	TSNSVRLG	612	0.440
KLRQVNAI	481	1.039	TVSSSPGT	403	0.434
NASMVANK	482	1.038	ANTNVTRP	847	0.428
VMSNVTQR	483	1.026	TPEPPTTA	929	0.426
GSGSVSKA	484	1.025	LSNNVRIK	424	0.426

[1628]

TAVKVDHS	485	1.025	TTAMHTVE	713	0.424
TSETPSTA	486	1.015	SKLVVSDR	585	0.423
STNTVRVN	487	1.009	LLNKVPLN	928	0.422
LSNQVSAR	488	1.005	SQPKVSTA	317	0.419
TTAPATGT	489	1.000	RKQAVSVA	887	0.419
TVLQKSTG	490	0.993	TVSDKSMS	379	0.415
SSSYVTVD	491	0.982	SGLNVTTR	798	0.415
TSVDKARV	492	0.978	TGSSPTGS	464	0.409
RRATVTTP	493	0.978	TREQTSTA	250	0.408
TGGRTPGG	494	0.971	MPNTVRAH	852	0.408
TDKQASGN	495	0.969	SSGDVKLW	277	0.406
TSPLANKS	496	0.963	VNNRVHQS	581	0.404
EGLKVLVG	497	0.949	SDNLVKTD	774	0.403
AIEKVSVN	498	0.943	TSPNRVSY	864	0.401
YVSNVAKA	499	0.942	LGNRVSST	807	0.399

[1629] 总之,这些结果表明,在NHP和大鼠中对具有环VIII修饰的此AAV5变体文库进行3轮筛选之后,多种AAV衣壳变体优于野生型AAV5,例如,在穿透血脑障壁(BBB)和脑中的表达方面。此外,鉴定出可能感染大鼠和NHP的衣壳变体,表明交叉物种趋向性和兼容性。

[1630] 实施例2.NHP、大鼠和小鼠中的个别衣壳表征

[1631] 此实施例描述在食蟹猕猴、挪威大鼠和BALB/c小鼠中进行静脉内注射后,相对于野生型AAV5,选自实施例1中所述的研究的AAV衣壳变体在中枢神经系统(CNS)中的转导水平、趋向性、穿过血脑障壁的能力和总体空间分布。所述衣壳变体为TTN-002(SEQ ID NO:982(氨基酸)和984(DNA),包含SEQ ID NO:943(由SEQ ID NO:944编码)),如上表3所概述。TTN-002的氨基酸和DNA序列分别提供于例如表4和表5中。

[1632] 用此衣壳变体囊封荧光素酶-EGFP转基因或由异源组成型启动子驱动的有效载荷-HA标签,生成AAV颗粒。通过静脉内以5e13个VG/kg向NHP(n=2)、以1e13个VG/大鼠(n=3;3e13个VG/kg)和/或以5e11个VG/小鼠(n=3;2e13个VG/kg)施用AAV颗粒制剂来测试每种衣壳变体及对照AAV5和/或AAV9衣壳。NHP和小鼠的生命周期为28天,且大鼠的生命周期为25天。接着收集各种CNS和外周组织来测量转基因mRNA、转基因蛋白和/或病毒DNA(生物分布)。施用于NHP和大鼠的AAV颗粒包含自身互补的病毒基因组,且施用于小鼠的AAV颗粒包

含单链病毒基因组。

[1633] A.NHP中的个别衣壳表征

[1634] 在静脉内施用包含TTN-002衣壳变体的AAV颗粒后28天,从NHP(食蟹猕猴)中分离出脑、脊髓和外周组织,包括心脏、肝和股四头肌,且通过qPCR来分析作为转基因表达的量度的转基因RNA的存在,且与AAV9对照进行比较。数据提供为转基因相对于管家基因的平均mRNA倍数变化以及相对于AAV9对照的倍数变化(表10)。如表10所示,相对于野生型AAV9对照,来自作为AAV5衣壳变体的TTN-002衣壳变体的mRNA转基因表达在NHP的脑中显著较高。更特定来说,在NHP的脑中,与野生型AAV9相比,来自TTN-002衣壳变体的mRNA表达大约高20-25倍。另外,在NHP的脊髓中,与野生型AAV9相比,来自TTN-002衣壳变体的mRNA表达大约高4-5倍。TTN-002还证实相对于AAV9对照,肝和DRG中的mRNA表达较低(表10)。

[1635] 还对从NHP中分离的脑、脊髓和外周组织分析作为病毒基因组水平的量度的病毒DNA的存在。表11中所提供的数据为每个二倍体基因组的平均DNA(病毒基因组(VG))拷贝数以及相对于AAV9对照的倍数变化。如表11所示,相对于野生型AAV9对照,AAV5衣壳变体TTN-002的生物分布在NHP脑中显著较高。相对于野生型AAV9对照,TTN-002在NHP肝中的生物分布较低。

[1636] 还对NHP的脑组织和脊髓进行免疫组织化学染色,以评估各个区域中的总体CNS趋向性和生物分布(图1A-1D)。免疫组织化学染色与qPCR分析相关,因为如与AAV9对照相比,TTN-002在脑(例如,整个大脑和小脑,图1A-1C)和脊髓(图1A和1D)中显示出显著较强的染色和有效载荷表达。更特定来说,TTN-002在颞叶皮层、鼻周皮层、苍白球、壳核、尾状核、丘脑、海马体、膝状体核、浦肯野层、小脑深部核团、齿状核、脑干和小脑中的神经元和神经胶质细胞中均证实定位、强有效载荷表达和转导(图1A-1C)。在齿状核中的星形胶质细胞中也观察到有效载荷表达。另外,在静脉内施用包含TTN-002衣壳变体的AAV颗粒后28天,对有效载荷和泛神经元标记物SMI-311的共表达的定量显示,在小脑中的脑深部核团的齿状核中的大约73.4%大神经元中存在有效载荷表达。在脊髓中,TTN-002在颈区(例如,C2)、胸区(例如,T10)和腰区(例如,L2)中证实定位、强有效载荷表达和转导(图1A和图1D)。另外,相对于野生型AAV9对照,TTN-002在DRG中显示出较少染色(大约少2倍)(图1D)。通过IHC测定发现,TTN-002和AAV9看来以相似的高效率转导肝和心脏(图1D)。另外,从NHP中分离的这些样品的组织病理学显示,在静脉内施用包含5e13个VG/kg的剂量的TTN-002衣壳变体和自身互补的病毒基因组的AAV颗粒后,NHP中无毒性迹象。

[1637] 表10.NHP中使用TTN-002衣壳变体时的转基因mRNA表达

[1638]

衣壳变体	组织	相对于管家基因的平均转基因mRNA倍数变化			相对于AAV9的倍数变化
		NHP1	NHP2	平均值	
AAV9	皮层	0.219	0.018	0.118	1.0
	丘脑	0.141	0.090	0.116	1.0
	小脑	0.677	0.098	0.388	1.0

	SC 颈部	0.610	2.040	1.325	1.0
	SC 胸部	0.796	1.238	1.017	1.0
	SC 腰部	4.950	0.890	2.920	1.0
	DRG 颈部	32.413	63.331	47.872	1.0
	DRG 胸部	27.509	57.666	42.588	1.0
	DRG 腰部	115.851	69.920	92.885	1.0
	肝脏	21.591	9.198	15.394	1.0
	心	231.804	20.067	125.936	1.0
	股四头肌	37.732	3.686	20.709	1.0
TTN-002	皮层	1.277	4.632	2.954	24.932
	丘脑	3.618	1.194	2.406	20.818
	小脑	3.839	14.762	9.301	23.998
	SC 颈部	3.683	9.095	6.389	4.822
	SC 胸部	5.759	6.639	6.199	6.096
	SC 腰部	7.544	16.550	12.047	4.126
	DRG 颈部	31.162	22.002	26.582	0.555
	DRG 胸部	25.228	14.915	20.071	0.471
	DRG 腰部	39.386	16.611	27.998	0.301
	肝	13.599	16.140	14.869	0.966
	心脏	380.487	143.202	261.844	2.079
	股四头肌	45.659	4.207	24.933	1.204

[1639]

[1640] 表11.NHP中使用TTN-002衣壳变体时的病毒DNA生物分布

衣壳变体	组织	每个二倍体基因组的平均 DNA 拷贝数			相对于 AAV9 的倍数变化
		NHP1	NHP2	平均值	
AAV9	皮层	0.213	0.074	0.1435	1.0
	丘脑	0.225	0.107	0.166	1.0
	小脑	0.199	0.079	0.139	1.0
	SC 颈部	0.688	0.103	0.396	1.0
	SC 胸部	0.385	0.192	0.288	1.0
	SC 腰部	0.214	0.21	0.212	1.0
	DRG 颈部	0.320	0.32	0.320	1.0
	DRG 胸部	0.303	0.733	0.518	1.0
	DRG 腰部	0.240	0.246	0.243	1.0
	肝	274.545	268.553	271.549	1.0
	心脏	2.367	2.833	2.600	1.0
	股四头肌	1.380	0.768	1.074	1.0
TTN-002	皮层	0.600	0.031	0.3155	2.203734
	丘脑	0.388	0.417	0.402	2.425078

[1641]

[1642]

	小脑	0.492	0.886	0.689	4.95212
	SC 颈部	0.008	0.298	0.153	0.386545
	SC 胸部	0.679	0.481	0.580	2.010803
	SC 腰部	0.801	0.194	0.497	2.347401
	DRG 颈部	0.711	0.636	0.674	2.106828
	DRG 胸部	0.000	0.305	0.153	0.294746
	DRG 腰部	0.356	0.517	0.437	1.795147
	肝	21.177	156.337	88.757	0.326855
	心脏	2.078	2.380	2.229	0.857412
	股四头肌	0.641	0.486	0.564	0.524939

[1643] B.大鼠中的个别衣壳表征

[1644] 在静脉内施用包含TTN-002衣壳变体(AAV5变体)的AAV颗粒后25天,从大鼠中分离脑和脊髓,且相对于野生型AAV5对照衣壳或野生型AAV9对照衣壳,分析作为转基因表达的量度的转基因RNA的存在(表12)。数据提供为转基因相对于管家基因的平均mRNA倍数变化以及相对于AAV5和AAV9对照的倍数变化(表12)。如表12所示,相对于AAV5和AAV9对照,来自TTN-002衣壳变体的mRNA转基因表达在脑和脊髓中较高。更特定来说,与野生型AAV5相比,来自TTN-002变体的转基因mRNA表达在大鼠脑和脊髓区(颈部、胸部和腰部)中大约高40-67倍,且与野生型AAV9相比,转基因mRNA表达在大鼠脑和脊髓区(颈部、胸部和腰部)中大约高5-7倍。

[1645] 对脑和脊髓组织及心脏外周组织进行免疫组织化学染色,以评估总体趋向性和生物分布。免疫组织化学染色与qPCR分析相关,因为TTN-002在大鼠的皮层、海马体、小脑和脊髓中显示相对于AAV9和AAV5增加的染色。TTN-002显示在大鼠的心脏中相对于AAV5增加的染色,但相对于AAV9减少的染色。

[1646] 表12.大鼠中使用TTN-002衣壳变体时的转基因mRNA表达

[1647]

组织	相对于管家基因的平均(n=3 只大鼠)转基因 mRNA 倍数变化			mRNA 表达相对于 AAV5 的倍数变化		mRNA 表达相对于 AAV9 的倍数变化	
	AAV5	AAV	TTN-	AAV9	TTN-002	AAV	TTN-00
		9	002			5	2
脑	0.051	0.381	2.051	7.533	40.510	0.133	5.378
脊髓 C1	0.121	1.045	8.209	8.629	67.796	0.116	7.856
脊髓 L1	0.118	1.647	4.901	13.947	41.499	0.072	2.975
脊髓 T1	0.133	1.066	6.952	8.015	52.291	0.125	6.524

[1648]

[1649] C.小鼠中的个别衣壳表征

[1650] 在静脉内施用包含TTN-002衣壳变体的AAV颗粒后28天,从BALB/c小鼠中分离出脑和肝,且通过qPCR来分析作为转基因表达的量度的转基因RNA的存在,且与AAV9和AAV5对照进行比较。数据提供为转基因相对于管家基因的平均mRNA倍数变化(表13)及转基因mRNA表达相对于AAV9和AAV5对照的倍数变化(表14)。如表13和表14所示,AAV5衣壳变体TTN-002证实在脑中相对于AAV9相似的转基因表达水平,及高于野生型AAV5的表达。如与野生型AAV5

相比,TTN-002衣壳变体在小鼠脑中的转基因mRNA表达高265.9倍(表14)。另外,如与野生型AAV9相比,野生型AAV5和AAV5衣壳变体TTN-002均导致肝中较低的转基因表达。

[1651] 还对从小鼠中分离的脑何肝分析作为病毒基因组水平的量度的病毒DNA的存在。表13中所提供的数据为每个二倍体基因组的平均DNA(病毒基因组(VG))拷贝数,且表14中所提供的数据为每个二倍体基因组的DNA拷贝数相对于AAV9和AAV5对照的倍数变化。AAV5衣壳变体TTN-002证实在小鼠脑中相对于AAV9可相当的生物分布,及与野生型AAV5相比增加的生物分布和病毒基因组水平。更特定来说,在脑中,TTN-002衣壳变体导致每个二倍体基因组的DNA(病毒基因组(VG))拷贝数相对于AAV5对照高9倍(表14)。此外,相对于AAV9,野生型AAV5和AAV5衣壳变体TTN-002导致肝中每个二倍体基因组的生物分布及DNA(病毒基因组(VG))拷贝数减少(表14)。

[1652] 表13.小鼠中使用TTN-002衣壳变体时的转基因mRNA表达

[1653]

组织	衣壳变体	小鼠	相对于管家基因的平均转基因 mRNA 倍数变化	每个二倍体基因组的平均转基因 DNA
脑	AAV5	小鼠 1	0.002	0.0152
		小鼠 2	0.004	0.0000
		小鼠 3	0.004	0.0114
		平均值	**0.003**	**0.0089**
	AAV9	小鼠 1	0.553	0.2000
		小鼠 2	0.235	0.1027
		小鼠 3	0.844	0.1213
		平均值	**0.544**	**0.1414**
	TTN-002	小鼠 1	0.179	0.1334
		小鼠 2	0.462	0.0555
		小鼠 3	1.828	0.0608
		平均值	**0.823**	**0.0832**
肝	AAV5	小鼠 1	0.395	0.9669
		小鼠 2	0.566	1.4158
		小鼠 3	0.882	1.4627
		平均值	**0.614**	**1.2818**
	AAV9	小鼠 1	9.164	34.1673
		小鼠 2	12.460	38.4415
		小鼠 3	5.190	24.0644
		平均值	**8.938**	**32.2244**
	TTN-002	小鼠 1	1.106	1.8292
		小鼠 2	1.305	1.6726
		小鼠 3	2.292	2.8985
		平均值	**1.568**	**2.1334**

[1654]

[1655] 表14.小鼠的脑和肝中转基因mRNA表达及每个二倍体基因组的DNA拷贝数相对于AAV9和AAV5的倍数变化

[1656]

组织	mRNA 表达相对于 AAV5 的倍数变化		mRNA 表达相对于 AAV9 的倍数变化	
	AAV9	TTN-002	AAV5	TTN-002
脑	175.796	265.874	0.006	1.512
肝	14.549	2.552	0.069	0.175

组织	每个二倍体基因组的 DNA 拷贝数相对于 AAV5 的变化倍数		每个二倍体基因组的 DNA 拷贝数相对于 AAV9 的变化倍数	
	AAV9	TTN-002	AAV5	TTN-002
脑	15.910	9.367	0.063	0.589
肝	25.140	1.664	0.040	0.066

[1657] 结论

[1658] 总之,这些数据表明,作为AAV5衣壳变体的TTN-002在NHP和啮齿动物(例如,大鼠和小鼠)中均为增强的CNS趋向性衣壳,且能够成功地穿透血脑障壁。更特定来说,TTN-002证实交叉物种的能力,从而证实在NHP和大鼠中增强的CNS趋向性和生物分布,以及在小鼠中相对于AAV5的改善。另外,TTN-002衣壳变体能够成功地穿透血脑障壁。

[1659] 实施例3.小鼠中TTN-002衣壳的成熟

[1660] 此实施例描述小鼠中TTN-002(SEQ ID NO:982(氨基酸)和984(DNA),包含SEQ ID NO:943)衣壳变体的成熟,以进一步增强其在中枢神经系统中的转导和生物分布,使AAV衣壳变体进一步演化,且提供交叉物种兼容性。使用两种方法使TTN-002衣壳序列成熟,以便在所述衣壳变体的环VIII内所包含的肽插入物内部及周围随机化和突变。在第一种成熟方法中,使用突变诱发引子以低频率引入点突变,这些点突变散布于TTN-002序列中的突变诱发区域中,介于大约位置571至位置586范围内,这些位置根据SEQ ID NO:982编号。在第二种成熟方法中,使三个连续氨基酸的集合在TTN-002序列中的突变诱发区域中随机化,其跨越大约位置571至位置586,这些位置根据SEQ ID NO:982编号。将由TTN-002的每种成熟方法产生的AAV衣壳变体汇集在一起,用于小鼠中的后续测试和表征。

[1661] 将使用第一种成熟方法由TTN-002生成的汇集的成熟AAV衣壳变体的文库及使用第二种成熟方法由TTN-002生成的汇集的成熟AAV衣壳变体的文库各自注射到三只CD-1远亲杂交小鼠中。在生活一段时间后,分离出小鼠的脑且提取RNA。在RNA回收和RT-PCR扩增之后,执行系统性NGS富集分析以计算相对于TTN-002对照的倍数富集比率,且鉴定出这些成熟变体内所包含的肽。来自第一种成熟方法的数据提供于表15中且来自第二种成熟方法的数据提供于表16。

[1662] 在来自第一种成熟方法的RNA回收和NGS分析后,基于其变异数系数(CV)来过滤成熟衣壳变体,所述CV是针对所采集的六个脑样品(每只小鼠两个)中的每种肽计算的。鉴定出CV值<1的那些,因为其为在自三只小鼠中分离的5/6或6/6脑样品中可靠检测到的肽。表15提供这些成熟衣壳变体的肽序列及所述成熟衣壳变体相对于未成熟TTN-002对照的倍数富集。如表15所示,大约28种TTN-002成熟衣壳变体证实相对于未成熟TTN-002对照的表达增加,其中大约16种变体证实至少2倍的表达增加。相对于未成熟TTN-002对照,数种变体证实至少8倍至15倍的表达增加。

[1663] 表15.第一种突变诱发方法后在远亲杂交小鼠的脑中的TTN-002成熟AAV衣壳变体

的NGS倍数富集

[1664]

序列	SEQ ID NO:	相对于TTN-002的倍数富集
TNSVSSYPAEVVQKTA	1537	15.404
TNNQSKYPAEVVQKTA	1538	10.261
TNNSSSYPAEVVQKTA	1539	8.539
TNNQSSYPATVVQKTA	1540	4.641
TNNSGSYPAEVVQKTA	1541	4.553
TNNLSKYPAEVVQKTA	1542	4.329
TNNQSSYPASVVQKTA	1543	3.849
TNNMSRYPAEVVQKTA	1544	3.586
TNNLSRYPAEVVQKTA	1545	3.580
TNNQSSYPPSIVQKTA	1546	3.529
TNNASKYPAEVVQKTA	1547	3.008
TNNQSSYPPSVVQKTA	1548	2.927
TNNLSSYPAEVVQKTA	1549	2.912
TNNQSSYPATLVQKTA	1550	2.688
TNSLSSYPAEVVQKTA	1551	2.243
TNQVSSYPAEVVQKTA	1552	2.157
TNNVSLYPAEVVQKTA	1553	2.075
TNNQSSYPVTVVQKTA	1554	1.857
TNNTSTYPAEVVQKTA	1555	1.753
TNNQSSLTAEVVQKTA	1556	1.612
TNNTSLYPAEVVQKTA	1232	1.484
TNNQSSYPAPVVQKTA	1558	1.478
TNNQSSYPASLIQKTA	1559	1.365
TNTISSYPAEVVQKTA	1560	1.263
TNNQSSYPAPLQQKTA	1561	1.261
TNNQSSYPAPLVQKTA	1562	1.149
TNNQSSYPPSLVQKTA	1300	1.141

[1665]

TNNPSRYPAEVVQKTA	1564	1.106
TNNQSSYPAEVVQKTA	1533	1.000
TNNSARYPAEVVQKTA	1566	0.980

[1666] 在来自第二种成熟方法的RNA回收和NGS分析后,在成熟衣壳变体中过滤出在注射成熟衣壳变体的所有小鼠的所有样品中可检测到的那些。表16提供这些成熟衣壳变体的肽序列,所述成熟衣壳变体相对于未成熟TTN-002对照的倍数富集,以及获得成熟变体的预测衣壳来源。如表16所示,大约526种TTN-002成熟衣壳变体证实相对于未成熟TTN-002对照的表达增加,其中大约358种变体证实至少2倍的表达增加。相对于未成熟TTN-002对照,数种变体证实20-711倍或更大的表达增加。

[1667] 表16.第二种突变诱发方法后在远亲杂交小鼠的脑中的TTN-002成熟AAV衣壳变体的NGS倍数富集

[1668]

序列	SEQ ID NO:	相对于TTN-002的倍数富集	序列	SEQ ID NO:	相对于TTN-002的倍数富集
TNNFFMYPAEVVQKTA	1000	711.864	TNNQSSYPAELVQKTA	1269	3.158
TNNQSSYPAEALLKTA	1001	107.670	TNSLWSYPAEVVQKTA	1270	3.150
TNNQGTYPAEVVQKTA	1002	87.226	TNNQSSYPAETVSKTA	1271	3.149
TISASSYPAEVVQKTA	1003	79.340	TNNISLYPAEVVQKTA	1272	3.114
TNGVHSYPAEVVQKTA	1004	37.131	TNNQSSYPAEVVQKST	1273	3.103
TNNHSAYPAEVVQKTA	1005	35.420	TNNQSSYPAELVSKTA	1274	3.082
TNNMAPYPAEVVQKTA	1006	34.417	TNNTGRYPAEVVQKTA	1275	3.067
TNNIATYPAEVVQKTA	1007	32.786	TYSTSSYPAEVVQKTA	1276	3.065
TNQISSYPAEVVQKTA	1008	30.799	TNTLSSYTAEVVQKTA	1277	3.040
TNNQSSYPAEVVQKLS	1009	30.539	TNSTSSYPAEVVQKTA	1278	3.033
TNNLGQYPAEVVQKTA	1010	28.520	TNAVFSYPAEVVQKTA	1279	3.025

[1669]

TNNPGAYPAEVVQKTA	1011	27.637	TNSTYSYPAEVVQKTA	1280	3.025
TNNQSSYPAEVVQKLK	1012	27.076	TNNLSMYPAEVVQKTA	1281	2.997
TNNQSSYPAAIIQKTA	1013	26.618	TNNQSSYPATLVQKPA	1282	2.973
TNNQSSYNTKVVQKTA	1014	22.549	TNNISVYPAEVVQKTA	1283	2.969
TNKITSYPAEVVQKTA	1015	22.308	TNNQSSYPAPVVQKTA	1284	2.937
TNNIGRYPAEVVQKTA	1016	22.159	TNTQSSYPAEVVQKTA	1285	2.929
TNLIASYPAEVVQKTA	1017	22.050	TNNQSSYPQSVVQKTA	1286	2.922
TNNTAGYPAEVVQKTA	1018	21.846	TNNSSVYPAEVVQKPA	1287	2.920
TNNMAKYPAEVVQKTA	1019	21.371	TNNQSSYPANQSQKTA	1288	2.889
TNNMAQYPAEVVQKTA	1020	20.203	TNNLAVYPAEVVQKTA	1289	2.888
TNNAGAYPAEVVQKTA	1021	19.342	TNNPSRYPAEVVQKTA	1290	2.877
TMSASSYPAEVVQKTA	1022	19.003	TNTTSSYPAEVVQKPA	1291	2.873
TNNAAVYPAEVVQKTA	1023	18.655	TNNQSSYPTNTVQKTA	1292	2.866
INNQSSYPAEVVQKTA	1024	18.426	TNNTGLYPAEVVQKTA	1293	2.862
TNNNSKYPAEVVQKTA	1025	18.030	TLLNSSYPAEVVQKTA	1294	2.847
TNNTSHYPAEVVQKTA	1026	17.574	TNNSALYPAEVVQKTA	1295	2.797
TNNLGSYPAEVVQKTA	1027	17.430	TNNLSLYPAEVVQKTA	1296	2.777
TNNQSSYPAEVISKTA	1028	17.172	TNNQSSYPAEVKSRTA	1297	2.760
TNLIDSYPAEVVQKTA	1029	16.804	TNNQSTATAEVVQKTA	1298	2.745
TNNTSGYPAEVVQKTA	1030	16.556	TNNQSSYPRDLVQKTA	1299	2.726
TNNQSSYPANDTQKTA	1031	15.769	TNNQSSYPPSLVQKTA	1300	2.719
TNNTAQYPAEVVQKTA	1032	15.607	TNSVSSYPAEVVQNTA	1301	2.703
TNLVSSYPAEVVQKTA	1033	15.329	TNNTAMYPAEVVQKTA	1302	2.700
TNNVSQYPAEVVQKTA	1034	15.269	TNNQGAYPAEVVQKTA	1303	2.678
TNNMSMYPAEV	1035	15.241	TNNQSSYPPTVV	1304	2.665

[1670]

VQKTA			QKTA		
TNSVSSYPAEV VQKTA	1036	14.066	TNNSAKYPAEV VQKTA	1305	2.645
TNSIGSYPAEVV QKTA	1037	13.549	TNNQSLYPAEV VQKTA	1306	2.630
TNNVSYYPAEV VQKTA	1038	13.380	TNNVSSYPAEVV QKPA	1307	2.622
TNNAASYPAEV VQKTA	1039	13.290	TNNLAPYPAEV VQKTA	1308	2.619
TNNMSIYPAEV VQKTA	1040	13.095	TNSPSSYPAEVV QKTA	1309	2.608
TNNVSGYPAEV VQKTA	1041	12.734	TNNHSSYPAEVV QKTA	1310	2.600
TNNMASYPAEV VQKTA	1042	12.673	TNNQSSYPAPVI QKTA	1311	2.599
TNNQSSYPAQIV QKTA	1043	12.669	TNNLSLYPAEVV QKPA	1312	2.598
TNKISSYPAEVV QKTA	1044	12.551	TNNVSKYPAEV VQKTA	1313	2.593
TNRVSSYPAEV VQKTA	1045	12.551	TNNQSYNPAEV VQKTA	1314	2.582
TNNVGYYPAEV VQKTA	1046	12.544	TNNPSVYPAEVV QKTA	1315	2.559
TNNLGAYPAEV VQKTA	1047	12.445	TNNIAPYPAEVV QKTA	1316	2.552
TNNTSQYPAEV VQKTA	1048	12.264	TNNQLTTPAEVV QKTA	1317	2.486
TNNQSSYPPKV VQKTA	1049	12.245	TNNLGKYPAEV VQKTA	1318	2.484
TNNVTMYPAEV VQKTA	1050	12.226	TNNQSSYPAHLI QKTA	1319	2.480
TNNQSSNPTEV VQKTA	1051	12.190	TNNMSTYPAEV VQKTA	1320	2.467
TNNIAWYPAEV VQKTA	1052	12.018	TNNQSSYPASVI QKTA	1321	2.457
TNNAAAYPAEV VQKTA	1053	11.885	TNNQSSYPAEVE SKLA	1322	2.455
TNNTASYPAEV VQKTA	1054	11.862	TNSVSSYPAEVV KKTA	1323	2.451
TNKVASYPAEV VQKTA	1055	11.641	TNNVSPYPAEVV QKTA	1324	2.437
TNNIGTYPAEVV QKTA	1056	11.597	TNNQSSYPPNM VQKTA	1325	2.413
TNNLSKYPAEV VQKTA	1057	11.521	TNNQSSYPANQT QKTA	1326	2.412
TNNLAKYPAEV VQKTA	1058	11.498	TNNQSRYPAEV VQKTA	1327	2.409
TNNQARYPAEV VQKTA	1059	11.473	TNTQASYPAEV VQKTA	1328	2.390

[1671]

TNNQSSYPATIIQKTA	1060	11.443	TNNQSSYPMPLVQKTA	1329	2.388
TNNISHYPAEVVQKTA	1061	11.290	TNNQSSYPREVVQKTA	1330	2.364
TKTTSSYPAEVVQKTA	1062	11.194	TNNSSSYPAEVVKKTA	1331	2.340
TNNQSSYPAEVVSKLA	1063	11.081	TNNQSSYPAPLQQKTA	1332	2.339
TNFVASYPAEVVQKTA	1064	11.061	TNLSGSYPAEVVQKTA	1333	2.326
TNNTSDYPAEVVQKTA	1065	11.043	TNNKSMYPAEVVQKTA	1334	2.319
TNNQSSYPPELVQKTA	1066	11.019	TNNQSSYPVTVVQKTA	1335	2.315
TNNMSLYPAEVVQKTA	1067	10.882	TNKQSSYPASVVQKTA	1336	2.312
TNNQEKYPAEVVQKTA	1068	10.869	TNNLSYYPAEVVQKTA	1337	2.311
TNNQSSYPADIVQKTA	1069	10.822	TNNSSLYPAEVVQKTA	1338	2.307
TNNMAMYPAEVVQKTA	1070	10.807	TNNQSSYPPNTVQKTA	1339	2.307
TNNQSSYPAELIEKTA	1071	10.660	TNNQSSYPALVIQKTA	1340	2.289
TNNQSSYPPDVVQKTA	1072	10.538	TNNQGVYPAEVVQKTA	1341	2.275
TNNLSAYPAEVVQKTA	1073	10.500	TNNQSRYPAEEVQKTA	1342	2.233
TNNIAHYPAEVVQKTA	1074	10.408	TNNQSSYPPHVVQKTA	1343	2.231
TNNHAKYPAEVVQKTA	1075	10.319	TNVVSSYPAEVVQKTA	1344	2.225
TNNMSRYPAEVVQKTA	1076	10.233	TNNQGTIPAEVVQKTA	1345	2.218
TNNTSSYPAEVVQKTA	1077	10.000	TNNQSSYPATLVQKTA	1346	2.200
TNNPAAYPAEVVQKTA	1078	9.842	TNNLSSYPAEVVQKTA	1347	2.190
TNSTFSYPAEVVQKTA	1079	9.812	TTLTSSYPAEVVQKTA	1348	2.129
TNNQSSYPAEVVQKTS	1080	9.758	TNLGGSYPAEVVQKTA	1349	2.125
TNNQSSYPAEVVQKSS	1081	9.758	TNNQSSYPAEVVNKLA	1350	2.107
TNNTSIYPAEVVQKTA	1082	9.739	TNNVGLYPAEVVQKTA	1351	2.106
TNNLAYYPAEVVQKTA	1083	9.674	TNNQSTPPAEVVQKTA	1352	2.094
TNNMGYYPAEV	1085	9.584	TNNSTIYPAEVV	1353	2.089

[1672]

VQKTA			QKTA		
TNLVSSYPAEV VQKPA	1086	9.577	TNNSAMYPAEV VQKTA	1354	2.073
TNNQAIYPAEV VQKTA	1087	9.556	TNTISSYPAEVV QKTA	1355	2.069
TNNPNSYPAEV VQKTA	1088	9.484	TNNQSSYPAALI QKTA	1356	2.069
TKTISSYPAEVV QKTA	1089	9.438	TNLGSSYPAEVV QKTA	1357	2.061
TNNIASYPAEVV QKTA	1090	9.423	TNKMSSYPAEV VQKTA	1358	2.056
TNMIASYPAEV VQKTA	1091	9.410	TNNQSSYPAEEV QKLK	1359	2.032
TNNMGAYPAEV VQKTA	1092	9.403	NNNQSSYPASV VQKTA	1360	2.032
TNAVHSYPAEV VQKTA	1093	9.332	TNNISPYPAEVV QKTA	1361	2.008
TNNISNYPAEVV QKTA	1094	9.304	TNNQSSYPPPVV QKTA	1362	2.000
TKSASSYPAEV VQKTA	1095	9.240	TNSVASYPAEVV QKTA	1363	1.983
TNNVGRYPAEV VQKTA	1096	9.183	TNNQSSYPAEVI RKTA	1364	1.974
TNNMSVYPAEV VQKTA	1097	9.158	TNNSSKYPAEVV QKTA	1365	1.972
TNNVSTYPAEV VQKTA	1098	9.126	TNNMALYPAEV VQKTA	1366	1.970
TNNTSAYPAEV VQKTA	1099	9.074	TNKLSSYPAEVV QKTA	1367	1.960
TNNQSSYPAEV VNKTA	1100	8.893	TNNQSSYPAPVV QKPA	1368	1.959
TNNQAMYPAEV VQKTA	1101	8.788	TNNQSSSVAEVV QKTA	1369	1.945
TNNQSSYPAEVI QKTA	1102	8.727	TNSIWSYPAEVV QKTA	1370	1.937
TNNASAYPAEV VQKTA	1103	8.711	TNNQRSYPATLV QKTA	1371	1.936
TNNLAAYPAEV VQKTA	1104	8.670	TNNQSSRDAEV VQKTA	1372	1.933
TNKVSSYPAEV VQKTA	1105	8.280	TNKVTSYPAEV VQKTA	1373	1.932
TNNQSSYPANLI QKTA	1106	8.247	TNNQAYYPAEV VQKTA	1374	1.928
TNNQSSYPSDIV QKTA	1107	8.206	TNNQSSYPATVV QKTA	1375	1.916
TNNQSSRESEVV QKTA	1108	8.198	TNNQSGISAEVV QKTA	1376	1.915
TNNISRYPAEVV QKTA	1109	8.130	TNNLALYPAEV VQKTA	1377	1.904

[1673]

TNNTAPYPAEVVQKTA	1110	8.050	TNNQAVYPAEVVQKTA	1378	1.889
TNNQSQYPAEVVQKTA	1111	7.947	TNNQSSNQAEVVQKTA	1379	1.864
TNNIGSYPAEVVQKTA	1112	7.845	TNNMSRYPAEVEQKTA	1380	1.857
TNNIGVYPAEVVQKTA	1113	7.821	TKTLSSYPAEVVQKTA	1381	1.855
TNNTNSYPAEVVQKTA	1114	7.794	TNNAWSYPAEVVQKTA	1382	1.848
TNNIARYPAEVVQKTA	1115	7.751	TNNQSSYPASMSQKTA	1384	1.832
TNNLSRYPAEVVQKTA	1116	7.723	TNLTPSYPAEVVQKTA	1385	1.828
TNNQSSYPEQVVQKTA	1117	7.705	TNLTSSYPAEVVQKTA	1386	1.818
TNKNASYPAEVVQKTA	1118	7.686	TNNLSSYPAEVEQKTA	1387	1.793
TNNTSTYPAEVVQKTA	1119	7.650	TNNQSNVTAEVVQKTA	1388	1.792
TNNSSIYPAEVVQKTA	1120	7.645	TNNQSSYPPQVVQKTA	1389	1.779
TNNQSSYPAEVVQKTT	1121	7.641	TNNQSSVSAEVVQKTA	1390	1.762
TNSGWSYPAEVVQKTA	1122	7.631	TNNQSSYPPNTVQKPA	1391	1.730
TNFIASYPAEVVQKTA	1123	7.624	TNLESSYPAEVVQKTA	1392	1.728
TNSTWSYPAEVVQKTA	1124	7.604	TNMVASYPAEVVQKTA	1394	1.715
TNNLGNYPAEVVQKTA	1125	7.585	TNNQSSTPSEVVQKTA	1395	1.709
TNNQSSYPPSIVQKTA	1126	7.576	TNNLSSYPAEVVQKPA	1396	1.698
TVASSSYPAEVVQKTA	1127	7.500	TNNQSPPRAEVVQKTA	1397	1.692
TNNQSSYPAELIAKTA	1128	7.498	TNSLGSYPAEVVQKTA	1398	1.680
TNNLARYPAEVVQKTA	1129	7.477	TNSTWSYPAEVVQKPA	1399	1.674
TNNQSSYPAEVVQKLN	1130	7.430	TNNQSSYPASVVQKPA	1400	1.672
TNNLSQYPAEVVQKTA	1131	7.420	TNNQSSYPAEVASKTA	1401	1.669
TNNISYYPAEVVQKTA	1132	7.337	TNNQSSYPAEVLHKTA	1402	1.664
TNNSSVYPAEVVQKTA	1133	7.278	TNTQSSYPASVVQKTA	1403	1.636
TNNQSSYPAEIV	1134	7.273	TNNQSSYPPSVE	1404	1.635

[1674]

HKTA			QKTA		
TNNLSVYPAEVVQKTA	1135	7.251	TNNQRSYPPSVVQKTA	1405	1.632
TNITGSYPAEVVQKTA	1136	7.246	TNNQSSYPAELVNKTA	1407	1.624
TNNMSKYPAEVVQKTA	1137	7.233	TNNLSGYPAEVVQKTA	1408	1.618
TNNCGSYPAEVVQKTA	1138	7.143	TNNQSSYPAEVVSKTA	1409	1.617
TNNVSSYPAEVVQKTA	1139	7.074	TNTVSSYTAEVVQKTA	1410	1.614
TNNIALYPAEVVQKTA	1140	7.043	TNSTSSYPAEVVQKPA	1411	1.614
TNNQLSIPAEVVQKTA	1141	7.017	TNNTSHYPAEEVQKTA	1412	1.608
TNNSSSYPAEVVQKPA	1142	6.889	TNNPSIYPAEVVQKTA	1413	1.605
TNNQSSYPAAVIQKTA	1143	6.870	TNTMSSYPAEVVQKTA	1414	1.602
TNNQSSYPAEIISKTA	1144	6.669	TNNQSHNPAEVVQKTA	1415	1.597
TNCVSSYPAEVVQKTA	1145	6.662	TNNQSSYPAELIHKTA	1416	1.592
TNNTARYPAEVVQKTA	1146	6.649	TNNQSSYPPSVVQKTA	1417	1.587
TNNQSSYPAEVINKTA	1147	6.627	TNNQSFYPAEVVQKTA	1418	1.574
TNSIASYPAEVVQKTA	1148	6.607	TNNSSSYPAEVVQQTA	1419	1.571
TNNQSSREAEVVQKTA	1149	6.499	TNTVSSYPAEVVQKTA	1420	1.570
TNNITRYPAEVVQKTA	1150	6.373	TNNQSPPSAEVVQKTA	1421	1.563
TNNTSVYPAEVVQKTA	1151	6.372	TNNQSTAPAEVVQKTA	1422	1.552
TNNTGSYPAEVVQKTA	1152	6.357	TNNQSSYPAEVVHKTA	1423	1.534
TNNMSQYPAEVVQKTA	1153	6.346	TNNQSSYPAMVVQKTA	1424	1.532
TNNISSYPAEVVQKTA	1154	6.335	TNNQATNPAEVVQKTA	1425	1.530
TNNSASYPAEVVQKTA	1155	6.319	TNNVSSYQAEVVQKTA	1426	1.530
TNSISSYPAEVVQKTA	1156	6.294	TNHVSSYPAEVVQKTA	1427	1.523
TNNISMYPAEVVQKTA	1157	6.235	TNNVSWYPAEVVQKTA	1428	1.520
TNNTATYPAEVVQKTA	1158	6.215	TNSLSSYPAEVVQKTA	1429	1.504

[1675]

TNNISTYPAEVVQKTA	1159	6.201	TNSVYSYPAEVVQKTA	1430	1.498
TNNVSHYPAEVVQKTA	1160	6.141	TNHTSSYPAEVVQKTA	1431	1.493
TNNSGSYPAEVVQKTA	1161	6.123	TNNQSSYPAPLVQKPA	1432	1.490
TNNLASYPAEVVQKTA	1162	6.091	TNSTTSYPAEVVQKTA	1433	1.483
TNNVSFYPAEVVQKTA	1163	6.088	TNNQSSYPAEVGSKLA	1434	1.482
TNNQSSYPVQVVQKTA	1164	6.075	TNNQSSYPVNTVQKTA	1435	1.477
TNNQAAYPAEVVQKTA	1165	6.059	TNNQSQSPAEVVQKTA	1436	1.471
TNQVSSYPAEVVQKTA	1166	6.058	TNNQSSYPPPLVQKTA	1437	1.468
TNSVGSYPAEVVQKTA	1167	6.020	TNNTSGYPAEVVQKPA	1438	1.463
TNTPASYPAEVVQKTA	1168	5.992	TNFSSSYPAEVVQKTA	1439	1.460
TNNLAMYPAEVVQKTA	1169	5.970	TNGVYSYPAEVVQKTA	1440	1.457
TNTTSSYPAEVVQKTA	1170	5.964	TNNQSSYPAPVVQQTA	1441	1.455
TNNSSSYPAEVVQKTA	1539	5.903	TNSLFSYPAEVVQKTA	1442	1.450
TNNQSSYPAEVVQNLK	1172	5.868	TNNQSSYPAEVIKKTA	1443	1.449
TNNLSIYPAEVVQKTA	1173	5.815	TNNQSSYPASMIQKTA	1444	1.446
TNNQSSYPAQVVQKTA	1174	5.773	TNNSVSYPAEVVQKTA	1445	1.445
TNTSASYPAEVVQKTA	1175	5.741	TNFVSSYPAEVVQKTA	1446	1.444
TNNTSKYPAEVVQKTA	1176	5.652	TNLISSYPAEVVQKTA	1447	1.427
TNNTTVYPAEVVQKTA	1177	5.625	TNNSAVYPAEVVQKTA	1448	1.418
TNNQSSYPPEVVQKTA	1178	5.541	TLSASSYPAEVVQKTA	1449	1.417
TNNQSSYPAEVMSRTA	1179	5.539	TNAVSSYPAEVVQKTA	1450	1.417
TNNASQYPAEVVQKTA	1180	5.532	TNNQSSTLIEVVQKTA	1451	1.411
TNNTSYYPAEVVQKTA	1181	5.525	TNNQSAYPAEVVQKTA	1452	1.404
TNNVSMYPAEVVQKTA	1182	5.522	TNNQSKYPAEVVHKTA	1453	1.402
TNNLGRYPAEV	1183	5.509	TNNLSIYPAEVV	1454	1.402

[1676]

VQKTA			QKPA		
TNNVAPYPAEV VQKTA	1184	5.492	TNPISSYPAEVV QKTA	1455	1.396
TNNQSIYPAEVV QKTA	1185	5.491	TNNQSSYPVAV VQKTA	1456	1.394
TNNQSKYPAEV VQKTA	1538	5.468	TNNSNMYPAEV VQKTA	1457	1.383
TNKIASYPAEVV QKTA	1187	5.445	TNNSNSYPAEVV QKTA	1458	1.379
TNNSSRYPAEV VQKTA	1188	5.231	TNNSSPYPAEVV QKTA	1459	1.362
TNNQSSYPAAV VQKTA	1189	5.228	SLVQSSYPAEVV QKTA	1460	1.358
TNNQSSYPAEV VQKLT	1190	5.112	TNNQSSYPAELK SKTA	1461	1.357
TNTIASYPAEVV QKTA	1191	5.112	TNSISSYTAEVV QKTA	1462	1.357
TNTLSSYPAEVV QKTA	1192	5.055	TNNQSSYPTNSV QKTA	1463	1.344
TNNQSSYPAKLI QKTA	1193	5.051	TNNSSQYPAEVV QKTA	1464	1.340
TNNQASYPAEV VQKTA	1194	5.004	TSLQSSYPAEVV QKTA	1465	1.339
TNNVGSYPAEV VQKTA	1195	4.953	TNNQSPYPAEVV QKTA	1466	1.325
HTRQSSYPAEV VQKTA	1196	4.804	TNNQSYYPASV VQKTA	1467	1.318
TNNMGSYPAEV VQKTA	1197	4.785	TNNLSAYTAEV VQKTA	1468	1.313
TNNVASYPAEV VQKTA	1198	4.765	TNNVSAYPAEV VQKTA	1469	1.307
TNNASGYPAEV VQKTA	1199	4.760	TNTVSSYPAEVV QKPA	1470	1.299
TNSPWSYPAEV VQKTA	1200	4.674	TNNNSVYPAEV VQKTA	1471	1.289
TNNVSNYPAEV VQKTA	1201	4.611	TNNQSSYPAPLV QKTA	1472	1.284
TNNQAKYPAEV VQKTA	1202	4.606	TNNQNSPPAEVV QKTA	1473	1.263
TNNQTTPPAEV VQKTA	1204	4.480	TNSVVSYPAEVV QKTA	1474	1.261
TSSASSYPAEVV QKTA	1205	4.465	TNNISKYPAEVV QKTA	1475	1.258
TNNTSFYPAEV VQKTA	1206	4.427	TNSPFSYPAEVV QKTA	1476	1.252
TNNLANYPAEV VQKTA	1207	4.415	TNNQSSYPAELL SKTA	1477	1.252
TNNTSRYPAEV VQKTA	1208	4.396	TNNQSSTRSEVV QKTA	1478	1.250

[1677]

TNNQSSYPASVVKKTA	1209	4.347	TNMVSSYPAEVVQKTA	1479	1.249
TLQQSSYPAEVVQKTA	1210	4.309	TNNQLPIPAEVVQKTA	1480	1.244
TNNQSSYPAGVIQKTA	1211	4.272	TNNTSMYPAEVVQKTA	1481	1.241
TNNQSSYPAETMSKTA	1212	4.266	TNNQSSYPSTVVQKTA	1482	1.240
TYTLSSYPAEVVQKTA	1213	4.256	TNNQSSYPATVVKKTA	1483	1.238
TNNSSSYTAEVVQKTA	1214	4.204	TNNQSSLTAEVVQKTA	1484	1.237
TNNTALYPAEVVQKTA	1215	4.200	TNNQSSYPPSVVQKPA	1485	1.224
TNNLTYYPAEVVQKTA	1216	4.198	TNNVTKYPAEVVQKTA	1486	1.210
TNNQSSYPAMVIQKTA	1217	4.192	TNNQSSYPPNVVQKTA	1487	1.210
TNNVTRYPAEVVQKTA	1218	4.189	TNNQSSYPAELMSKTA	1488	1.205
TNNMSAYPAEVVQKTA	1219	4.186	TNALSSYPAEVVQKTA	1489	1.202
TNNQSSYPAEVIAKTA	1220	4.186	TNNQSSYPANVVQKTA	1490	1.201
TNNLAQYPAEVVQKTA	1221	4.176	TNNSALYPAEVVQKPA	1491	1.201
TNNTGGYPAEVVQKTA	1222	4.111	TVTQSSYPAEVVQKTA	1492	1.197
TNNQSSYPAEIVQKTA	1223	4.093	TNNQSSVTLEVVQKTA	1493	1.191
TNNISAYPAEVVQKTA	1224	4.078	TNNVSVYPAEVVQKTA	1494	1.190
TNNQSMYPAEVVQKTA	1225	4.058	TNNQSSYPQELVQKTA	1495	1.188
TNTLGSYPAEVVQKTA	1226	4.043	TNNQSSYPASVNQKTA	1496	1.180
TNNQSSYPAESVQKTA	1227	4.042	TNNQSSYPASIIQKTA	1497	1.171
TNNPSGYPAEVVQKTA	1228	3.980	TNNQSSYPAPVVKKTA	1498	1.163
TNNQDNMPAEVVQKTA	1229	3.917	TNNLAFYPAEVVQKTA	1499	1.158
TNNVSLYPAEVVQKTA	1230	3.868	TNNQSSYPAEAVSKTA	1500	1.157
TNNASSYPAEVVQKTA	1231	3.832	TNNQSSYPAPLEQKTA	1501	1.156
TNNTSLYPAEVVQKTA	1232	3.828	TNNLSKYPAEVVQKTD	1502	1.154
TNNQSSYPASV	1233	3.809	TNNQSSYPAPLI	1503	1.143

[1678]

VQKTA			QKTA		
TNNQSSYPAELISKTA	1234	3.772	TNTLASYPAEVVQKTA	1504	1.140
TNSVHSYPAEVVQKTA	1235	3.770	TNNQSSYPPGVVQKTA	1505	1.139
TNSVWSYPAEVVQKTA	1236	3.762	TNNQSTQRAEVVQKTA	1506	1.125
TTTPSSYPAEVVQKTA	1237	3.732	TNNLSRYTAEVVQKTA	1507	1.122
TTTTSSYPAEVVQKTA	1238	3.728	TNTVSSYPAEEVQKTA	1508	1.122
TNNQSSYPESNVQKTA	1240	3.683	TNNQSSVAAEVVQKTA	1509	1.106
TNNQSSYPPDLVQKTA	1241	3.672	TNNHSRYPAEVVQKTA	1510	1.101
TNLVASYPAEVVQKTA	1242	3.621	TNNQSSYPPAVVQKTA	1511	1.088
TNLVSSYQAEVVQKTA	1243	3.575	TNNVSRYPAEVVQKTA	1512	1.078
TNNQSVYPAEVVQKTA	1244	3.568	TNKQSSYPAPVVQKTA	1513	1.074
TNNQGYYPAEVVQKTA	1245	3.550	TNNQSSYPAEVSSKTA	1514	1.072
TNSPRSYPAEVVQKTA	1248	3.508	TNNSSHYPAEVVQKTA	1515	1.072
TNNQSSYPASLIQKTA	1249	3.504	TNNSPSYPAEVVQKTA	1516	1.070
TNNSTRYPAEVVQKTA	1250	3.500	TNNQSSYPSTIVQKTA	1517	1.070
TNNSGAYPAEVVQKTA	1251	3.497	TNATSSYPAEVVQKTA	1518	1.069
TNNLAIYPAEVVQKTA	1252	3.450	TNVSLSYPAEVVQKTA	1519	1.068
TNHVASYPAEVVQKTA	1253	3.436	TNNQSSYPASMVQKTA	1520	1.067
TNNQSKYPAEVEQKTA	1254	3.434	TNLVSSYPAEVLQKTA	1521	1.058
TNNHASYPAEVVQKTA	1255	3.413	TNNQSSLNAEVVQKTA	1522	1.057
TNSESSYPAEVVQKTA	1256	3.392	TNNPSKYPAEVVQKTA	1523	1.050
TNNQSSYPANSVQKTA	1257	3.344	TNNQSSYPASIVQKTA	1524	1.046
TNNQSSYPAELVHKTA	1258	3.344	TNNTVSYPAEVVQKTA	1525	1.045
TNNQSSYPPNSVQKTA	1259	3.313	TNNVALYPAEVVQKTA	1526	1.045
TNNTAKYPAEVVQKTA	1260	3.304	TNNQSSYPPVVVQKTA	1527	1.014

[1679]

TNNKTSYPAEVVQKTA	1261	3.271	TNNQSSYPSSVVKKTA	1528	1.014
TNNNSNYPAEVVQKTA	1262	3.220	TNNQSSYPAEVLKKTA	1529	1.011
TNTLSSYQAEVVQKTA	1263	3.212	TNNQSTLTAEVVQKTA	1530	1.011
TNRISSYPAEVVQKTA	1264	3.210	TNNQSSYPAPLVQQTA	1531	1.005
TNNQSSYPATLIQKTA	1265	3.200	TNNLSPYPAEVVQKTA	1532	1.001
TNNGVSYPAEVVQKTA	1266	3.191	TNNQSSYPAEVVQKTA	1533	1.000
TNNTSSYPAEVVQKPA	1267	3.173	TNNQLTSPAEVVQKTA	1534	0.995
TNNQSSYPAEVVQKTP	1268	3.165	TNNLSHYPAEVVQKTA	1535	0.991

[1680] 这些数据表明,在两种成熟方法后,与相应未成熟衣壳变体相比,在小鼠中生成具有显著增强的CNS趋向性的具有环VIII修饰的成熟TTN-002衣壳变体,其已在小鼠、大鼠和/或NHP的脑中展现相对于AAV5和/或AAV9的显著倍数富集。

[1681] 实施例4.NHP中TTN-002衣壳的成熟

[1682] 此实例描述NHP、具体地食蟹猕猴中TTN-002(SEQ ID NO:982(氨基酸)和984(DNA),包含SEQ ID NO:943(由SEQ ID NO:944编码))衣壳变体的成熟,以进一步增强其在中枢神经系统和外周组织中的转导和生物分布,且使AAV衣壳变体进一步演化。使用两种方法使TTN-002衣壳序列成熟,以便在所述衣壳变体的环VIII内所包含的肽插入物内部和周围随机化和突变。在第一种成熟方法中,使用突变诱发引子以低频率引入点突变,这些点突变散布于TTN-002序列中的突变诱发区域中,介于大约位置571至位置586范围内,这些位置根据SEQ ID NO:982编号。在第二种成熟方法中,使三个连续氨基酸的集合在TTN-002序列中的突变诱发区域中随机化,其跨越大约位置571至位置586,这些位置根据SEQ ID NO:982编号。将由TTN-002的每种成熟方法产生的AAV衣壳变体汇集在一起,用于NHP中之后续测试和表征。

[1683] 将使用第一种成熟方法由TTN-002生成的汇集的成熟AAV衣壳变体的文库及使用第二种成熟方法由TTN-002生成的汇集的成熟AAV衣壳变体的文库各自注射到两只NHP中。在生活一段时间后,分离出NHP的脑、心脏、肝、肌肉和DRG且提取RNA。在RNA回收和RT-PCR扩增之后,执行系统性NGS富集分析以计算相对于TTN-002对照的倍数富集比率,且鉴定出这些变体内所包含的肽。

[1684] 在来自第二种成熟方法的RNA回收和NGS分析后,基于其变异数系数(CV)来过滤成熟衣壳变体,CV是针对自两只NHP采集的脑、心脏、肝、肌肉和DRG样品中的每种肽计算的。鉴定出CV值<2的那些,因为其为在从两只NHP的脑中分离的大多数样品中可靠检测到的肽。

[1685] 表17提供这些成熟衣壳变体的肽序列及所述成熟衣壳变体相对于未成熟TTN-002对照的倍数富集,这证实在第一种成熟方法和第二种成熟方法后,NHP的脑中的表达相对于未成熟TTN-002衣壳变体的最大倍数变化。如表17所示,在第一种成熟方法后,大约5种TTN-002成熟衣壳变体证实相对于未成熟TTN-002对照的表达增加,这证实NHP脑中的表达相对

于未成熟TTN-002对照的至少5倍至53倍增加。在第二种成熟方法后,大约38种TTN-002成熟衣壳变体证实NHP脑中的表达相对于未成熟TTN-002对照的增加,其中至少27种证实至少2倍的表达增加(表17)。数种变体证实NHP脑中的表达相对于未成熟TTN-002对照的至少12倍至222倍增加(表17)。

[1686] 在每种成熟方法后,还针对NHP的DRG、肌肉、肝(RNA和DNA)和心脏计算表17中的TTN-002成熟变体的表达的倍数变化。在NHP的脑中导致相对于未成熟TTN-002变体的表达增加的数种变体还导致其他组织中的表达增加。例如,相对于未成熟TTN-002对照,包含氨基酸序列TNNQSSYTPSLVQKTA(SEQ ID NO:1585)的成熟TTN-002衣壳变体在脑、心脏和肝中证实增加的表达。同样,相对于未成熟TTN-002对照,包含氨基酸序列TNNQSSYPPSLVKKTA(SEQ ID NO:1591)和TNNQSSYPPSLVQKPA(SEQ ID NO:1593)的成熟TTN-002衣壳变体在脑和心脏中证实增加的表达。另外,相对于未成熟TTN-002对照,包含氨基酸序列INNQSSYPAEVVQKTA(SEQ ID NO:1024)的成熟TTN-002衣壳变体在脑和肌肉中证实增加的表达。又,如表17所示,多种在脑中具有增加的表达的TTN-002衣壳变体在DRG中脱靶。因此,数种成熟变体在NHP的超过一种组织类型中证实增加的趋向性,其中多种在DRG中显示减少的表达。

[1687] 表17.如与其他NHP和小鼠组织相比,在第一种和第二种突变诱发方法后,NHP的脑中TTN-002成熟AAV衣壳变体的NGS倍数富集

[1688]

序列	SEQ ID NO:	相对于 TTN-002 的倍数富集						
		脑 (NHP)	DRG (NHP)	肝 RNA (NHP)	肝 DNA (NHP)	心脏 (NHP)	肌肉 (NHP)	脑 (小鼠)
第一种成熟方法								
TNNQSKYPAEVVQKTA	1538	53.751	0.000	2.270	0.518	0.000	0.000	10.261
TNNTSLYPAEVVQKTA	1232	22.903	0.000	2.824	0.000	0.000	0.000	1.484
TNNSSSYPAEVVQKTA	1539	19.002	0.000	1.240	0.000	0.000	0.000	8.539
TNNQSRYPAEVVQKTA	1327	8.499	0.000	0.846	0.516	0.000	0.000	1.704
TNNQSSYPPSLVQKTA	1300	5.458	0.000	0.966	0.000	5.100	0.000	1.141
TNNQSSYPAEVVQKTA	1533	1.000	1.000	1.000	1.000	1.000	1.000	1.000
第二种成熟方法								
TNNAGAYPAEVVQKTA	1021	222.785	0.000	0.000	0.000	0.000	0.000	19.342
TNNIGSYPAEVVQKTA	1112	93.586	0.000	0.000	0.142	0.000	0.000	7.845
TNTASSYPAEVVQKTA	1575	46.343	0.000	0.019	2.353	0.234	0.000	0.012
TNNTSLYPAEVVQKTA	1232	45.861	0.000	5.118	0.432	0.059	0.087	3.828
TNNLGSYPAEVVQKTA	1027	44.023	0.000	0.000	0.156	0.000	0.000	17.430
TNNQSTNKAE	157	42.63	0.000	0.268	4.032	0.000	0.000	0.000

[1689]

VVQKTA	8	3						
TNNHSSYPAEVVQKTA	1310	38.584	0.000	0.275	0.430	0.076	0.000	2.600
TNNQSKYPAEVVQKTA	1538	20.857	0.000	2.180	1.568	0.001	0.000	5.468
TNNSSSYTAEVVQKTA	1214	19.625	0.000	0.000	2.152	0.000	0.000	4.204
TNNQSKYPAEVEQKTA	1254	15.501	0.000	3.630	0.000	0.000	0.000	3.434
TNNTSLYPAEEVQKTA	1583	14.603	0.000	1.276	0.000	0.000	0.000	1.249
TNTASSYQAEVVQKTA	1584	13.180	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSYTPSLVQKTA	1585	12.552	0.000	5.700	2.591	24.396	0.000	1.831
TNNQSSYPPSLVQKTA	1300	9.123	0.000	2.178	1.093	20.059	0.052	2.719
TNNQSRYPAEVVQKTA	1327	8.503	0.000	1.887	1.125	0.002	0.000	2.409
TNNSSSYPAEVVQKTA	1539	8.092	0.267	0.854	0.904	0.003	0.000	5.903
TNNQSRYPAEEVQKTA	1342	7.267	0.000	0.879	0.783	0.000	0.000	2.233
TNNQSSYPPSLEQKTA	1590	6.924	0.000	1.616	0.000	8.962	0.000	1.550
TNNQSSYPPSLVKKTA	1591	6.625	0.000	0.000	0.891	10.953	0.000	1.783
TNNLSSYQAEVVQKTA	1592	5.913	276.643	0.000	0.000	0.000	0.000	1.280
TNNQSSYPPSLVQKPA	1593	5.789	0.000	2.977	1.161	25.392	0.000	3.704
TNNSSSYPAEVVKKTA	1331	4.191	0.000	0.766	0.521	0.000	0.000	2.340
TNNQSKYPAEVVHKTA	1453	3.201	0.000	0.766	0.000	0.000	0.000	1.402
TNNSSSYPAEVVQKPA	1142	2.748	0.000	0.794	0.000	0.000	0.000	6.889
INNQSSYPAEVVQKTA	1024	2.472	0.000	0.992	0.695	0.275	75.541	18.426
TNNHSSYPAAVVQKTA	1598	2.007	0.000	0.000	0.000	0.000	0.000	0.587
TNSQSSNPAEVVQKTA	1599	1.829	0.000	0.000	0.000	0.000	0.000	0.116
TNNSSSYPAEVVQQTA	1419	1.645	0.000	0.510	0.000	0.000	0.000	1.571
NNNQSRYPAEVVQKTA	1601	1.513	0.000	0.565	0.000	0.000	0.000	0.480
TNNQSSYTAEVVQKNA	1602	1.350	0.000	0.000	0.891	0.000	0.000	0.159

[1690]

TNNTSLCPAEVVQKTA	1603	1.307	0.000	0.851	0.000	0.000	0.000	0.132
TNNQRSYTAEVVQKTA	1604	1.186	0.000	2.836	2.348	0.000	0.000	0.407
TNSTSLYPAEVVQKTA	1605	1.186	0.000	0.000	0.000	0.000	0.000	0.196
TNNQSSYTAEVVKKTA	1606	1.143	0.000	0.839	2.325	0.000	0.000	0.697
TNNHSSYPAEVLQKTA	1607	1.136	0.000	0.000	0.000	0.000	0.000	0.163
TNNQSSYQAEEVQKTA	1608	1.113	0.000	0.000	0.000	0.000	0.000	0.561
TNNQSSYPAEVVQKTA	1533	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TNNKSKYPAEVVQKTA	1610	0.980	0.000	0.000	2.953	0.000	0.000	0.600
TNNQSRYPAEVMQKTA	1611	0.967	0.000	0.000	0.000	0.000	0.000	0.085
TNSTSPYPAEVVQKTA	1612	0.880	0.000	0.000	0.000	0.000	0.000	0.211
TNNQSSYPAWLIQKTA	1613	0.713	0.000	0.000	0.000	0.000	0.000	0.091
TNNQSSYPAEATKKTA	1614	0.693	0.000	0.408	6.529	0.000	0.000	0.000
TNKQSSYTAEVVQKTA	1615	0.678	0.000	0.000	1.151	0.000	0.000	0.240
TNKIGSYPAEVVQKTA	1616	0.660	0.000	0.000	0.812	0.000	0.000	0.046
TNNKSRYPAEVVQKTA	1617	0.648	0.000	5.889	2.080	3.490	0.000	0.467
TNNQSSYPAEVVQKTD	1618	0.646	0.000	0.756	1.247	0.945	3.270	0.530
TNNQSSYPAEVVQKNA	1619	0.640	4.200	3.327	3.647	0.622	3.655	0.505
TNNQRKYPAEVVQKTA	1620	0.629	0.000	6.158	2.709	0.000	0.000	0.171
TNNQSSNPAEEVQKTA	1621	0.617	0.000	0.000	0.989	0.000	0.000	0.449
NNNQSSYPAEVVQKTA	1622	0.546	0.000	1.197	1.528	1.025	58.686	0.750
TNNQSYYPAEEVQKTA	1623	0.537	0.000	0.000	0.000	0.000	0.000	0.365
TKNHSSYPAEVVQKTA	1624	0.522	0.000	0.000	0.845	0.000	0.000	0.062
TNNQASYPAEVVQKTA	1586	88.651						
TNKQASYPAEVVQKTA	1587	1.100						

[1691] 此外,如表17所示在NHP的脑中在第一种和第二种成熟方法后证实相对于未成熟TTN-002对照的表达增加的数种TTN-002成熟衣壳变体也在小鼠中在第一种和第二种成熟

方法后证实小鼠的脑中的表达增加。例如,包含氨基酸序列TNNQSKYPAEVVQKTA(SEQ ID NO:1538)的成熟TTN-002衣壳变体在NHP的脑中在第一种成熟方法后证实相对于未成熟TTN-002的53.7倍表达增加,在NHP的脑中在第二种成熟方法后证实相对于未成熟TTN-002的20.86倍表达增加,在小鼠的脑中在第一种成熟方法后证实相对于未成熟TTN-002的10.26倍表达增加,且在小鼠的脑中在第二种成熟方法后证实相对于未成熟TTN-002的5.47倍表达增加。同样,包含氨基酸序列TNNSSSYPAEVVQKTA(SEQ ID NO:1539)的成熟TTN-002衣壳变体在NHP的脑中在第一种成熟方法后证实相对于未成熟TTN-002的18.997倍表达增加,在NHP的脑中在第二种成熟方法后证实相对于未成熟TTN-002的8.093倍表达增加,在小鼠的脑中在第一种成熟方法后证实相对于未成熟TTN-002的8.539倍表达增加,且在小鼠的脑中在第二种成熟方法后证实相对于未成熟TTN-002的5.903倍表达增加。相对于未成熟TTN-002对照,包含SEQ ID NO:1021、1024、1027、1112、1142、1214、1232、1254、1300、1310、1327、1331、1342和1593的氨基酸序列的成熟TTN-002衣壳变体也证实NHP和小鼠的脑中的表达增加。因此,在至少两种不同物种中,数种成熟变体证实相对于未成熟对照的表达增加,表明交叉物种趋向性。

[1692] 表18提供这些成熟衣壳变体的肽序列及所述成熟衣壳变体相对于未成熟TTN-002对照的倍数富集,这证实在第二种成熟方法后NHP的心脏中的表达增加。如表18所示,相对于未成熟TTN-002对照,大约17种TTN-002成熟衣壳变体证实心脏中的表达增加,其中至少13种证实至少2倍的表达增加。数种变体证实心脏中的表达相对于未成熟TTN-002对照的至少10倍至47倍增加。针对NHP的脑、DRG、肌肉和肝(RNA和DNA)及小鼠的脑计算表18中的TTN-002成熟变体的表达的倍数变化。

[1693] 表18.如与其他NHP和小鼠组织相比,在第二种突变诱发方法后,NHP的心脏中TTN-002成熟AAV衣壳变体的NGS倍数富集

[1694]

序列	SEQ ID NO	相对于 TTN-002 的倍数富集						
		心脏 (NHP)	脑 (NHP)	DRG (NHP)	肝 RNA (NHP)	肝 DNA (NHP)	肌肉 (NHP)	脑 (小鼠)
TNSSLSYQAEVVQKTA	2064	47.380	0.000	0.000	0.000	0.000	0.000	0.000
TNSSLSYTAEVVQKTA	2065	45.827	0.066	0.000	4.885	8.710	0.000	0.000
TNSSLYYPAEVVQKTA	2066	41.710	0.000	0.000	0.000	0.000	0.000	0.000
TNTSATYPAEVVQKTA	2067	35.136	0.000	0.000	0.000	0.756	0.000	0.476
TNSSLSYPAEVVQKTA	2068	32.805	0.015	0.000	0.744	2.122	0.000	0.002
TNSSLSYPAEEVQKTA	2069	23.357	0.000	0.000	0.000	0.000	0.000	0.000
TNSSLSYPAEVVQKTD	2070	18.293	0.000	0.000	0.000	0.000	0.000	0.000
TNSSLSYPAEVVQNTA	2071	17.081	0.000	0.000	0.000	0.000	0.000	0.000
TNSSLSYPAEVVQKNA	2072	13.129	0.000	0.000	0.000	0.000	0.000	0.000
TNSSLSYPAEVVQKPA	2073	13.115	0.000	0.000	0.000	0.521	0.000	0.000
TNNQTSTEAEVVQKTA	2074	10.056	0.000	0.000	0.000	0.000	0.000	0.000
TNKSATYPAEVVQKTA	2075	9.090	0.000	0.000	0.000	0.000	0.000	0.000
TNSSLSYPADVVQKTA	2076	4.692	0.000	0.000	0.000	0.384	0.000	0.000
TNNQSSYPAEVVQKTA	1533	1.000	1.000	1.000	1.000	1.000	1.000	1.000
TNNQSSYPAEVVQKNA	2078	0.622	0.640	4.200	3.327	3.647	3.655	0.505
TNNQSSYPSTDVQKTA	2079	0.572	0.020	0.000	0.000	3.216	0.000	0.000
TNNKGLYPAEVVQKTA	2080	0.506	0.000	0.000	0.000	5.411	0.000	0.002

[1695] 表19提供这些成熟衣壳变体的肽序列及所述成熟衣壳变体相对于未成熟TTN-002对照的倍数富集,这证实在第二种成熟方法后NHP的肌肉中的表达增加。如表19所示,相对于未成熟TTN-002对照,大约33种TTN-002成熟衣壳变体证实肌肉中的表达增加,其中至少19种证实至少2倍的表达增加。数种变体证实肌肉中的表达相对于未成熟TTN-002对照的至少7倍至38倍增加。针对NHP的脑、DRG、心脏和肝(RNA和DNA)及小鼠的脑计算表19中的TTN-002成熟变体的表达的倍数变化。

[1696] 表19.第二种突变诱发方法后在NHP的肌肉中的TTN-002成熟AAV衣壳变体的NGS倍

数富集

[1697]

序列	SEQ ID NO	相对于 TTN-002 的倍数富集						
		肌肉 (NHP)	脑 (NHP)	DRG (NHP)	肝 RNA (NHP)	肝 DNA (NHP)	心脏 (NHP)	脑 (小鼠)
TNNQSSYDFTVVQKTA	2024	38.128	0.000	267911.1	0.000	4.209	0.032	0.000
TNNQSSYPAEVVQNVG	2025	20.373	0.000	82501.777	7.770	0.000	0.015	0.000
TNNQSPHAAEVVQKTA	2026	9.977	0.000	85036.503	2.478	0.841	0.009	0.003
TNNLHVYPAEVVQKTA	2027	9.969	0.000	59488.293	0.000	0.934	0.006	0.000
TLKTSSYPAEVVQKTA	2028	9.646	0.000	69692.245	0.032	1.906	0.155	0.000
TNNQAASPAEVVQKTA	2029	7.483	0.000	147630.5	0.000	1.848	0.023	0.000
TNLSMSYPAEVVQKTA	2030	5.960	0.000	67462.229	0.220	1.697	0.014	0.005
TKAFSSYPAEVVQKTA	2031	5.743	0.000	42730.728	0.000	0.919	0.000	0.000
TAYQSSYPAEVVQKTA	2032	4.703	0.000	61779.917	0.022	1.908	0.000	0.000
TNNQSQHQAEVVQKTA	2033	4.510	127.1	107992.3	3.186	4.569	0.015	0.000
TNNQSSVQLEVVQKTA	2034	3.846	0.000	54250.016	7.862	4.173	0.013	0.000
TNNQSSAFAEVVQKTA	2035	3.503	32.731	67253.425	3.207	5.348	0.004	0.000
TNNQSAVLAEVVQKTA	2036	3.351	0.000	28296.228	2.069	3.784	0.007	0.000
LNAQSSYPAEVVQKTA	2037	3.311	0.000	76314.798	0.000	0.476	0.000	0.000
TNNQSQSSAEVVQKTA	2038	3.310	0.000	198372.9	2.579	0.766	0.007	0.003

[1698]

TNNQSSLHAEVVQKTA	2039	2.685	0.013	23403.089	4.165	4.578	0.005	0.001
TNNFQLYPAEVVQKTA	2040	2.460	0.000	28791.699	0.000	0.053	0.000	0.000
TNNQSSYPAEVVQSMR	2041	2.099	0.000	176004.5	0.000	0.000	0.000	0.000
TLLYSSYPAEVVQKTA	2042	2.044	0.000	8688.765	0.000	2.362	0.006	0.000
TNNQSSTLDEVVQKTA	2043	1.947	0.000	0.000	0.000	0.043	397.2	0.001
TNNQTTTPAEVVQKTA	2044	1.839	0.000	59308.623	1.658	2.863	0.003	0.004
TNNQSSYPAEVVELLA	2045	1.800	0.000	12934.481	0.000	0.175	0.004	0.000
TNNQVNYPAEVVQKTA	2046	1.694	0.000	0.000	12.363	2.400	0.000	0.000
TTSTSSYPAEVVQKTA	2047	1.576	0.002	42212.609	8.264	2.618	0.000	0.003
TNNQSSYPQIAVQKTA	2048	1.506	0.000	0.000	28.933	3.364	0.000	0.000
TNNQSSYPAENLPKTA	2049	1.327	0.000	56068.064	0.000	2.002	0.000	0.008
TNNQSSYPASSPQKTA	2050	1.304	0.000	0.000	1.637	1.814	466.4	0.000
TNNQFMVPAEVVQKTA	2051	1.151	0.000	63928.318	0.000	2.409	0.000	0.000
TNNTTTYPAEVVQKTA	2052	1.121	0.073	12941.562	0.719	0.808	0.065	0.008
TPSKSSYPAEVVQKTA	2053	1.083	0.000	1.366	7.344	1.632	0.000	0.000
TNYLMSYPAEVVQKTA	2054	1.078	0.000	24113.689	0.000	1.271	0.382	0.000
TNLAMSYPAEVVQKTA	2055	1.074	0.000	77469.008	0.701	1.184	0.000	0.006
TNNQSSLMQEVVQKTA	2056	1.057	0.000	0.000	21.968	3.563	0.000	0.000
TNNQSSYPNAIVQKTA	2057	0.960	0.000	0.000	0.077	0.929	0.019	0.000
TNNQSSYPAEVRTATA	2058	0.915	0.000	0.000	30.158	1.431	0.000	0.000
TNNQSSSMTEVVQKTA	2059	0.788	66.451	33379.460	1.383	3.930	0.000	0.000
TNLSPSYPAEVVQKTA	2060	0.778	0.008	11925.049	2.716	1.300	0.000	0.000
TNNQSSYPAESSRKTA	2061	0.657	0.000	0.000	7.755	3.289	377.3	0.000
TNNPLTYPAEVVQKTA	2062	0.641	0.000	18890.829	0.000	1.060	0.000	0.000
TNNLSTYPAE	206	0.595	0.029	32495.4	0.000	1.113	0.065	0.01

[1699]

VVQKTA	3			46				3

[1700] 表20提供这些成熟衣壳变体的肽序列及所述成熟衣壳变体相对于未成熟TTN-002

对照的倍数富集,这证实在第一种和第二种成熟方法后NHP的肝中的表达增加。如表20所示,在第一种成熟方法后,大约7种TTN-002成熟衣壳变体证实肝中的表达相对于未成熟TTN-002对照的11倍至189倍增加。在第二种成熟方法后,大约395种TTN-002成熟衣壳变体证实肝中的表达相对于未成熟TTN-002对照的至少9倍增加(表20)。数种变体证实肝中的表达相对于未成熟TTN-002对照的至少50倍至114倍增加(表20)。针对NHP的脑、DRG、心脏和肌肉及小鼠的脑计算表20中的TTN-002成熟变体的表达的倍数变化。

[1701] 表20.两种突变诱发方法后在NHP的肝中的TTN-002成熟AAV衣壳变体的NGS倍数富集

[1702]

序列	SEQ ID NO	相对于 TTN-002 的倍数富集						
		肝 RNA (NHP)	肝 DNA (NHP)	脑 (NHP)	DRG (NHP)	心脏 (NHP)	肌肉 (NHP)	脑 (小鼠)
第一种成熟方法								
TDYHRGDP AKVVQKSA	1625	189.324	0.000	0.000	0.000	0.000	0.000	0.000
TNKQKFSLT EELQKNA	1626	101.263	5.032	0.000	0.000	0.000	0.000	0.000
GKTPRHFTQ QVVQENA	1627	84.647	1.661	0.000	0.000	0.000	0.000	0.000
SENERSDTS VVVAQIL	1628	56.534	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSIRL EVVQKTA	1629	13.829	2.357	0.000	0.000	0.000	0.000	0.000
TNNQSSSLA EVVQKTA	1630	11.832	4.625	29.220	0.000	0.000	0.000	0.000
TNNQSGPTA EVVQKTA	1631	11.391	0.000	0.000	0.000	0.000	0.000	1.678
第二种成熟方法								
TNNQSSYPA EAPKKTA	1632	114.219	5.207	419.144	0.000	0.000	0.000	0.000
TNNQSSYPR KVVQKTA	1633	104.291	0.000	0.000	0.000	0.000	0.000	0.045

[1703]

TNNQSRIQA EVVQKTA	1634	94.646	4.673	0.000	0.000	0.000	0.182	0.001
TNNQSWHQ AEVVQKTA	1635	90.698	8.068	0.000	0.000	0.000	0.000	0.000
TNNQSGPTA EVVQKTA	1631	78.597	5.825	0.000	0.000	0.005	0.000	5.645
NEWQSSYP AEVVQKTA	1636	73.960	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSIKWA EVVQKTA	1637	73.487	1.668	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVVNIQA	1638	71.285	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA HNSQKTA	1639	71.153	1.854	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVVQYQD	1640	67.017	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSRQQ EVVQKTA	1641	63.063	7.092	0.002	0.000	0.000	0.660	0.000
TNNQNCSPA EVVQKTA	1642	59.544	0.951	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ERSHKTA	1643	53.889	4.190	0.000	0.000	0.000	0.000	0.004
TVNRSSYPA EVVQKTA	1644	53.711	6.381	0.000	0.000	0.000	0.000	0.000
TNNQHGGP AEVVQKTA	1645	53.380	0.565	0.000	0.000	0.000	0.000	0.000
TNNQSMIW AEVVQKTA	1646	52.362	1.443	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EASRKTA	1647	49.543	3.268	0.000	0.000	0.040	0.000	0.000
TNNQSSYPA HLTQKTA	1648	49.237	4.942	0.000	0.000	0.000	0.000	0.000
TNNRRGYP AEVVQKTA	1649	48.153	0.000	0.000	0.000	0.000	0.000	0.000
TNNKRIYPA EVVQKTA	1650	48.033	0.000	0.000	0.000	0.000	0.000	0.000
TNNQGALP AEVVQKTA	1651	47.761	1.178	0.000	0.000	0.000	0.000	0.036
TNNQSSYPA EASKKTA	1652	44.984	6.761	0.010	0.000	0.000	0.000	0.000
TGQHSSYPA EVVQKTA	1653	43.610	2.355	0.000	0.000	0.000	0.000	35.110
TNNQSSIVQ EVVQKTA	1654	43.603	5.228	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ERNTKTA	1655	43.470	5.264	0.000	0.000	0.007	0.000	0.000
TNNQSSYPA QALQKTA	1656	43.207	2.932	0.000	0.000	0.000	0.000	0.000
TNNQSRRG	1657	42.38	2.20	0.000	0.00	0.00	0.000	0.00

[1704]

AEVVQKTA		4	5		0	0		0
TNNQSSYPQVMVQKTA	1658	42.160	0.019	0.000	0.000	0.000	0.000	0.000
TNNQSSQKREVVQKTA	1659	42.065	5.253	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAENARKTA	1660	42.025	4.024	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAGTGQKTA	1661	41.566	3.734	0.000	0.000	0.000	0.000	0.000
TRMPSSYPAEVVQKTA	1662	40.545	2.585	0.000	0.000	0.000	0.000	0.002
TNNQSSPISEVVQKTA	1663	39.647	1.460	0.000	0.000	0.000	0.000	0.000
TNNQSSYGAVVVQKTA	1664	39.126	1.251	0.000	0.000	0.000	0.385	0.000
TNNQSSYPAEQPVKTA	1665	38.694	1.536	0.000	0.000	0.000	0.000	0.000
TNNQSSGYREVVQKTA	1666	37.888	3.088	0.000	0.000	0.003	0.000	0.000
TNNQSGAQAEVVQKTA	1667	37.624	5.498	0.000	0.000	0.000	0.000	0.000
TNNQSSKGREVVQKTA	1668	37.172	2.348	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAESKHKTA	1669	36.198	1.594	0.000	0.000	0.000	0.000	0.000
TNNQSSYNMRVVQKTA	1670	35.863	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAEKNTKTA	1671	34.039	2.322	0.000	0.000	0.000	0.000	0.083
TNNQSSYPAERREKTA	1672	33.598	12.673	0.000	0.000	0.000	0.000	0.000
TNNQWWMPAEVVQKTA	1673	33.472	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSYTHAEVVQKTA	1674	33.128	4.433	0.000	0.000	0.000	0.000	0.004
TNNAAKYPAEVVQKTA	1675	32.786	3.029	0.009	0.000	0.000	0.330	0.007
TNNQSSYPLVCVQKTA	1676	32.659	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSNAYAEVVQKTA	1677	32.331	0.978	0.000	0.000	0.000	0.000	0.000
TGVKSSYPAEVVQKTA	1678	32.102	0.884	0.000	0.000	0.000	0.000	0.000
TNNQSSKKTEVVQKTA	1679	31.801	10.223	0.000	0.000	0.012	0.000	0.000
TNNQSSYPAEMANKTA	1680	31.613	3.467	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA	1681	30.87	3.13	0.002	0.00	0.00	0.000	0.00

[1705]

SMGQKTA		5	1		0	0		0
TNNQSSYPA PGIQKTA	1682	30.648	1.520	0.000	0.000	0.000	0.000	0.015
TNNQSSYPA QDKQKTA	1683	30.569	0.444	0.000	0.000	0.000	0.000	0.000
TNNQSSYPE RQVQKTA	1684	30.436	6.881	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA SQSQKPA	1685	30.284	0.810	0.000	0.000	0.000	0.000	0.000
TNNQSQGT AEVVQKTA	1686	30.072	2.154	0.002	0.000	0.000	0.000	0.014
TNNQPRLPA EVVQKTA	1687	29.628	2.410	0.000	0.000	0.014	0.000	0.003
TNNQSSYPA ELGKKTA	1688	29.441	5.305	0.000	0.000	0.000	0.000	0.009
TNNQSSLTK EVVQKTA	1689	29.297	2.313	0.000	0.000	0.000	0.000	0.001
ELCQSSYPA EVVQKTA	1690	29.051	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSYPV SRVQKTA	1691	28.960	1.001	0.000	0.000	0.000	0.000	0.000
TNNQSSHHQ EVVQKTA	1692	28.825	3.732	0.000	0.000	0.000	0.000	0.000
TNNQWCSP AEVVQKTA	1693	28.739	0.354	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVVQREY	1694	28.521	0.000	0.000	0.000	0.000	0.000	0.000
THKSSSYPA EVVQKTA	1695	28.280	1.162	0.000	0.000	0.000	0.000	0.010
TGNVSSYPA EVVQKTA	1696	28.028	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSYPT SDVQKTA	1697	27.923	3.549	0.000	0.000	0.130	0.479	0.009
TNNQSSYPA INTQKTA	1698	27.590	5.013	0.000	0.000	0.000	0.000	0.000
TNAPRSYPA EVVQKTA	1699	27.525	2.321	0.000	0.000	0.000	0.000	0.002
TNNQAHNP AEVVQKTA	1700	27.459	2.355	0.000	0.000	0.000	0.000	0.000
TNNQSARIA EVVQKTA	1701	27.305	5.056	0.000	0.000	0.000	0.000	0.000
TNNQSSIHQ EVVQKTA	1702	27.261	3.709	0.000	0.000	0.000	0.000	0.011
TNNQSSYPG PKVQKTA	1703	26.890	4.998	0.000	0.000	0.000	0.000	12.243
TNNQSSYPA SQSQKTA	1704	26.085	3.225	0.000	0.000	0.001	0.000	0.000
TNNQSVAK AEVVQKTA	1705	25.379	2.456	0.004	0.000	0.000	0.000	0.000

[1706]

TGLLSSYPA EVVQKTA	1706	25.248	1.073	0.000	17.422	0.000	0.000	0.866
TNNQSSYPA EMNKKTA	1707	25.179	11.154	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVVIRGA	1708	24.885	0.454	0.000	0.000	0.000	0.000	0.000
TNNQSSYVL EVVQKTA	1709	24.760	3.119	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ERSAKTA	1710	24.694	4.869	0.001	0.000	0.000	0.000	0.000
TNNQSSYRT NVVQKTA	1711	24.382	0.619	0.000	0.000	0.000	0.000	0.000
TNFTPSYPA EVVQKTA	1712	24.209	0.528	0.000	1.558	0.000	0.000	0.005
TNNQSSTTT EVVQKTA	1713	24.072	0.872	0.005	0.000	0.058	0.000	0.001
TNNQSFLDA EVVQKTA	1714	24.054	0.885	0.008	0.000	0.000	0.000	0.006
TNNQSSYPT SNVQKTA	1715	23.530	3.586	0.000	0.000	0.000	0.622	0.000
TNNQSSRHD EVVQKTA	1716	23.512	1.200	0.000	0.000	0.000	0.000	0.000
TNNQSSYPQ KSVQKTA	1717	23.414	2.234	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ERHPKTA	1718	23.276	4.332	0.000	0.000	0.000	0.000	0.000
TNNQSSKVQ EVVQKTA	1719	23.231	2.813	0.000	0.000	0.000	0.000	0.000
LGVQSSYPA EVVQKTA	1720	23.175	2.605	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EARQKTA	1721	22.889	4.006	0.000	0.000	0.000	0.000	0.000
TNNQSSYPD SRVQKTA	1722	22.614	4.720	0.000	0.000	0.000	0.000	0.000
TNNQSSYPQ MDVQKTA	1723	22.445	2.469	0.000	0.000	0.000	0.000	0.000
TNNQSVQG AEVVQKTA	1724	22.409	0.161	0.000	0.000	0.000	0.000	0.000
TNNQSSYSE AVVQKTA	1725	22.354	0.911	0.000	0.000	0.000	1.593	0.000
TNNQSSYTA LVSQKTA	1726	22.345	8.610	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVRAKTA	1727	22.283	5.669	0.013	0.000	0.000	0.000	0.002
TNNQSSPGT EVVQKTA	1728	22.240	5.707	0.000	0.000	0.000	0.000	0.000
TNNQSMMK AEVVQKTA	1729	22.043	2.468	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA	1730	21.89	0.25	0.000	0.00	0.00	0.000	0.07

[1707]

DNNQKTA		9	4		0	0		4
TNNQSSQNLEVVQKTA	1731	21.884	4.064	0.000	0.000	0.000	0.000	0.000
TNNKPPYPAEVVQKTA	1732	21.774	1.524	0.000	0.000	0.000	0.000	0.002
TNNQSNTYAEVVQKTA	1733	21.680	4.281	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAERPGKTA	1734	21.642	4.857	0.000	0.000	0.015	0.000	0.002
TNNQSSYPAESVRKTA	1735	21.641	5.644	0.000	0.000	0.000	0.784	0.002
TNNQSSYSVGVVQKTA	1736	21.595	2.004	0.000	0.000	0.000	0.000	0.000
TNNQSSYPLDRVQKTA	1737	21.582	4.140	0.000	0.000	0.005	0.000	0.000
TNNQSQIKAEVVQKTA	1738	21.553	5.308	0.000	0.000	0.000	0.000	0.000
TNNQSSNMVEVVQKTA	1739	21.435	2.898	0.000	0.000	0.000	0.000	0.000
TNNQSSYSSHVVQKTA	1740	21.301	2.342	0.000	0.000	0.000	0.000	0.000
TNNQGSSPAEVVQKTA	1741	21.255	2.814	0.000	0.000	0.000	0.000	0.006
TNNQSSKYSEVVQKTA	1742	21.233	2.717	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAEARAKTA	1743	21.084	3.135	0.000	0.000	0.000	0.000	0.000
TNNQSSYTASLTQKTA	1744	21.040	14.386	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAEGPQKTA	1745	21.024	1.251	0.000	0.000	0.000	0.000	0.000
TNNQSQPKAEVVQKTA	1746	20.892	1.972	0.000	1.841	0.000	0.000	0.000
TNNQSQVIAEVVQKTA	1747	20.571	0.344	0.000	0.000	0.000	0.000	0.007
TNNQSSYPAERPNKTA	1748	20.524	5.620	0.000	0.000	0.000	0.000	0.000
TNNQSWTVAEVVQKTA	1749	20.134	3.768	0.000	0.000	0.000	0.000	0.000
TNNQSRVIAEVVQKTA	1750	20.128	2.855	0.000	0.000	0.000	0.506	0.000
TNNQSSYPAEYERKTA	1751	19.648	1.217	0.000	0.000	0.000	0.000	0.000
TFHKSSYPAEVVQKTA	1752	19.633	8.119	0.000	0.000	0.000	0.000	0.000
TNNQRERPAEVVQKTA	1753	19.581	2.717	0.000	0.000	0.000	0.000	0.022
TNNQSSYPAERQGKTA	1754	19.516	6.304	0.000	0.000	0.000	0.000	0.000

[1708]

TNNQSSYPA EMTKKTA	1755	19.43 4	2.04 7	0.000	0.00 0	0.00 0	0.000	0.01 6
TNNQSWNM AEVVQKTA	1756	19.40 3	5.71 8	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSLKD EVVQKTA	1757	19.29 7	1.73 4	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA GPSQKTA	1758	18.94 1	3.24 4	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSKVN EVVQKTA	1759	18.82 2	1.30 9	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPM RDVQKTA	1760	18.82 1	3.80 4	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA LTNQKTA	1761	18.74 0	2.25 9	0.000	0.00 0	0.00 0	0.000	0.00 1
TNNQSSNAL EVVQKTA	1762	18.62 3	3.78 4	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA EVVQKSQ	1763	18.24 0	2.24 5	0.000	0.00 0	0.00 0	0.000	0.00 0
TPYKSSYPA EVVQKTA	1764	18.23 8	0.79 8	0.000	0.00 0	0.00 0	0.000	0.00 0
TTTPSSYPA EVVQKTA	1765	18.19 6	2.66 9	0.007	1.39 2	0.00 0	0.000	3.73 2
TNNQSSYKD AVVQKTA	1766	18.13 1	6.01 4	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA EDTRKTA	1767	17.93 3	5.00 9	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA ERTQKTA	1768	17.91 2	4.90 7	0.000	0.00 0	0.00 3	0.000	0.00 0
TNNQSSYPA EHAQKTA	1769	17.80 0	0.92 9	0.000	0.00 0	0.00 0	0.000	0.01 1
TNNQSSYPA ASPQKTA	1770	17.78 3	2.26 1	0.000	0.00 0	0.08 3	0.000	0.00 0
TNNQSSKNR EVVQKTA	1771	17.77 2	7.50 1	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA EVTTKTA	1772	17.47 4	2.14 8	0.000	0.00 0	0.00 0	0.000	0.00 4
TNNQSSYPA EVVSKQA	1773	17.43 9	2.94 5	0.000	0.00 0	0.00 0	0.000	0.04 1
TNNQRNQP AEVVQKTA	1774	17.35 2	5.41 6	0.000	0.00 0	0.00 0	0.000	0.00 5
TNNQSSYPA EKVSKTA	1775	17.30 0	3.16 1	0.000	0.00 0	0.00 0	0.000	0.00 8
TNNQSSYPA EFTKKTA	1776	17.28 7	4.12 3	0.000	0.00 0	0.00 0	0.000	0.00 0
TAMDSSYPA EVVQKTA	1777	17.27 8	6.71 4	0.000	0.00 0	0.00 0	0.000	0.00 0
YLKQSSYPA EVVQKTA	1778	17.18 5	2.03 2	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA	1779	17.17	4.71	0.005	5.52	0.00	0.000	0.00

[1709]

EKLTKTA		4	4		7	0		0
TNNQSSYPA EVVNGKA	1780	17.143	1.029	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EKGSKTA	1781	17.038	5.983	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA HINQKTA	1782	16.961	7.399	0.000	0.000	0.000	0.000	0.000
TNNQSSYPV QNVQKTA	1783	16.911	3.355	0.000	0.000	0.000	0.000	0.000
TNNQSSYPK DTVQKTA	1784	16.890	1.297	0.000	0.000	0.000	0.000	0.000
TNNQVIWPA EVVQKTA	1785	16.816	2.555	0.000	0.000	0.000	0.000	0.000
TNNQSWVA AEVVQKTA	1786	16.775	0.627	0.000	0.000	0.000	0.000	0.000
TDGISSYPA EVVQKTA	1787	16.708	0.150	0.000	0.000	0.000	0.000	0.000
TNNQSSRNG EVVQKTA	1788	16.667	0.367	0.000	0.000	0.000	0.000	0.000
TNNQSSHPA EVVQKTA	1789	16.653	3.700	6.125	0.000	0.006	0.000	0.897
TNNQSSMQ MEVVQKTA	1790	16.616	6.837	0.000	0.000	0.000	0.000	0.000
TNNQSSVSH EVVQKTA	1791	16.522	4.597	0.000	0.000	0.000	12.027	0.004
TNNQSSTFT EVVQKTA	1792	16.482	2.661	0.000	0.000	0.000	0.000	0.000
TNNQSSYTA TASQKTA	1793	16.475	5.560	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ESPKKTA	1794	16.426	3.157	0.023	0.000	0.000	0.000	0.000
TNNQSSYPA EPRGKTA	1795	16.249	6.345	0.000	0.000	0.000	0.000	0.000
TNNQSSIRL EVVQKTA	1796	16.197	3.887	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA KDLQKTA	1797	16.197	0.255	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EHVSKTA	1798	16.196	5.228	0.000	0.000	0.000	0.000	0.032
TNNQSSYPA EGKAKTA	1799	16.075	3.013	0.000	0.000	0.000	0.000	0.026
TNNQSRMK AEVVQKTA	1800	15.914	1.084	0.000	0.000	1.069	0.000	0.000
TNNQSWLQ AEVVQKTA	1801	15.911	3.287	0.000	0.000	0.009	0.000	0.000
TNNQSHDR AEVVQKTA	1802	15.847	3.669	0.000	0.000	0.000	0.000	0.000
TNNQSSGST EVVQKTA	1803	15.778	5.237	0.000	0.000	0.000	0.000	0.000

[1710]

TNNQSSYPA EPAKKTA	1804	15.763	3.448	0.000	0.000	0.000	0.000	0.005
TNNQSSYPA ERSNKTA	1805	15.735	3.609	0.000	0.000	0.000	0.000	0.000
TFNKSSYPA EVVQKTA	1806	15.700	0.476	0.000	0.000	0.000	0.000	0.000
TNNQSCLYA EVVQKTA	1807	15.679	2.156	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ERVLKTA	1808	15.674	3.436	0.000	0.000	0.000	0.000	0.000
TFRRSSYPA EVVQKTA	1809	15.621	10.989	0.000	0.000	0.000	0.000	0.000
TNNRHNYP AEVVQKTA	1810	15.620	1.758	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA ETEGKTA	1811	15.607	1.076	0.000	0.000	0.000	0.000	0.000
TNHPSSYPA EVVQKTA	1812	15.593	1.681	0.000	0.000	0.000	0.000	0.002
TTGISSYPAE VVQKTA	1813	15.532	1.301	0.000	0.000	0.000	0.000	0.000
TNAAQSYPA EVVQKTA	1814	15.531	1.512	0.000	0.000	0.000	0.000	0.000
TNNQSSGSM EVVQKTA	1815	15.459	5.283	0.000	0.000	0.000	0.000	0.000
TNNARGYP AEVVQKTA	1816	15.450	4.233	0.000	0.000	0.000	0.000	0.000
TNNQSSYPF RSVQKTA	1817	15.429	0.000	0.000	0.000	0.000	0.000	0.000
TNLTGSYPA EVVQKTA	1818	15.427	1.420	0.000	0.000	0.000	0.000	0.001
TNNQSSTYL EVVQKTA	1819	15.377	2.983	0.004	0.000	0.000	0.000	0.000
TNNQSSYPP AQVQKTA	1820	15.332	3.214	0.000	0.000	0.005	0.000	0.000
THYGSSYPA EVVQKTA	1821	15.294	0.251	0.000	0.000	0.000	0.000	0.000
TNNQSRFGA EVVQKTA	1822	15.279	2.490	0.000	0.000	0.000	0.000	0.000
TNNVKHYP AEVVQKTA	1823	15.239	0.208	0.000	0.000	0.000	0.000	0.000
TNNQSSYPT KSVQKTA	1824	15.221	3.402	0.010	0.000	0.000	0.000	0.047
TNNQSSYSQ HVVQKTA	1825	15.207	2.302	0.000	0.000	0.000	0.000	0.003
TNNQSSGHT EVVQKTA	1826	15.202	7.933	0.000	0.000	0.000	0.000	0.005
TNNQSSYPA QTVQKTA	1827	15.159	3.586	0.000	0.000	0.077	0.000	0.002
TNNQSSYPA	1828	15.14	1.61	0.000	0.00	0.00	0.000	0.00

[1711]

HFPQKTA		5	0		0	0		0
TNNQSSYPLAKVQKTA	1829	15.138	3.309	0.000	0.000	0.000	0.000	0.000
TNNQSSYTALTVQKTA	1830	15.057	10.012	0.000	0.000	0.000	0.000	0.000
TNQRSSYPAEVVQKTA	1831	15.029	3.694	0.000	0.000	0.000	0.000	1.493
TNNQSSYPAEVVSKKA	1832	14.986	1.378	0.000	0.000	0.000	0.000	0.003
TNNQSSYPASRSQKTA	1833	14.966	1.741	0.000	0.000	0.000	0.000	0.005
TNNQSDIRAEVVQKTA	1834	14.861	9.186	0.000	0.000	0.000	0.000	0.005
TNNQYSRPAEVVQKTA	1835	14.818	6.222	0.000	0.000	0.000	0.000	0.672
TNNQSGLNAEVVQKTA	1836	14.713	4.785	0.005	0.000	0.000	0.000	0.000
TNNQSRMMAEVVQKTA	1837	14.700	4.917	0.000	0.000	113.3	0.387	0.000
TSAKSSYPAEVVQKTA	1838	14.673	1.418	0.000	0.000	0.000	0.000	0.009
TNNQSSYPANGPQKTA	1839	14.649	2.642	0.000	0.000	0.000	0.000	0.000
VFYQSSYPAEVVQKTA	1840	14.578	0.183	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAERTSKTA	1841	14.574	3.973	0.000	0.000	0.002	0.000	0.000
TNSTPSYPAEVVQKTA	1842	14.527	2.179	0.072	0.000	0.000	0.000	1.229
TNNQSSYPAEHAVKTA	1843	14.486	2.272	0.000	0.000	0.000	0.000	0.000
TNNQSRGHAEVVQKTA	1844	14.451	3.993	0.000	0.000	0.000	0.185	0.000
TNNRAGYPAEVVQKTA	1845	14.424	4.286	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAEQRLKTA	1846	14.420	6.535	0.000	0.000	0.002	0.000	0.000
TNNQSRIMAEVVQKTA	1847	14.362	4.485	0.000	0.000	0.086	0.000	0.000
TNNQSTTLAEVVQKTA	1848	14.334	4.654	0.000	0.000	0.004	0.109	0.004
TNNQSSYPAEVANKTA	1849	14.333	5.878	0.000	0.000	0.000	0.000	0.004
TNNQSSSVKEVVQKTA	1850	14.316	3.803	0.019	0.000	0.000	0.000	0.000
TNNQSSYTTSLVQKTA	1851	14.263	9.339	0.000	0.000	0.000	0.000	0.053
TNNQSSYPAHNMQKTA	1852	14.168	2.792	0.000	0.000	0.000	0.000	0.000

[1712]

TNNQSSYPA QTTQKTA	1853	14.11 2	3.44 4	0.000	0.00 0	0.00 0	0.000	0.00 3
TNNQSSYPF NSVQKTA	1854	14.02 7	1.74 7	0.001	0.00 0	0.00 0	0.000	0.00 2
TNNQSIIRAE VVQKTA	1855	13.96 2	3.87 2	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSLRIA EVVQKTA	1856	13.95 8	2.10 5	0.001	0.00 0	0.00 0	0.000	0.00 0
TNNQSCCLA EVVQKTA	1857	13.93 8	3.96 7	0.000	0.00 0	0.00 0	1.090	0.00 0
TNNQSSALL EVVQKTA	1858	13.92 6	4.60 7	0.017	0.00 0	0.00 0	0.000	0.00 1
TNNQSSYPA EVVQRNT	1859	13.90 1	0.85 9	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQLNLPA EVVQKTA	1860	13.88 6	1.55 7	0.000	0.00 0	0.00 0	0.000	0.00 3
TNNQSRMY AEVVQKTA	1861	13.87 2	4.62 6	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA EVVLDMA	1862	13.84 7	0.02 3	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSKVM AEVVQKTA	1863	13.83 6	4.27 0	0.000	0.00 0	0.00 6	0.000	0.00 0
TNNQSSVRQ EVVQKTA	1864	13.78 2	5.31 1	0.032	0.00 0	0.00 5	22413 0.7	0.00 1
TNNQSSYVS KVVQKTA	1865	13.77 8	5.00 7	0.000	0.00 0	0.00 0	1.049	0.00 0
TNNQSSYPA ETHVKTA	1866	13.75 0	4.73 7	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSHGI EVVQKTA	1867	13.71 2	4.71 7	0.000	0.00 0	0.00 0	0.000	0.00 0
THSQSSYPA EVVQKTA	1868	13.67 0	1.73 4	0.001	0.00 0	0.00 0	0.000	2.20 2
TNNQSSYPG NAVQKTA	1869	13.65 8	3.40 8	0.000	0.00 0	0.00 0	0.000	0.00 1
TNNQSSMN LEVVQKTA	1870	13.62 7	3.09 4	0.000	0.00 0	0.00 0	0.000	0.00 1
TNNQSSYTA RVVQKTA	1871	13.62 4	1.41 8	0.039	0.00 0	0.00 0	0.000	0.00 0
TNNQSQGA AEVVQKTA	1872	13.58 2	4.66 9	0.000	0.00 0	0.00 0	0.000	0.00 9
TNNQSSYPA ERGAKTA	1873	13.57 1	4.71 2	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSPILE VVQKTA	1874	13.55 7	6.00 2	0.000	0.00 0	0.00 0	0.000	0.00 5
TNNQSSYPA ETHSKTA	1875	13.53 0	6.58 6	0.000	0.00 0	0.00 0	0.000	0.00 0
TAFRSSYPA EVVQKTA	1876	13.52 5	2.36 1	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNVRSYPA	1877	13.46	3.99	0.000	0.00	0.00	0.000	0.01

[1713]

EVVQKTA		5	3		0	3		0
TNNQSATPAEVVQKTA	1878	13.413	2.222	0.000	0.270	0.000	0.000	0.003
TNNQSLTYAEVVQKTA	1879	13.296	1.688	0.000	0.000	0.000	0.000	0.000
TNNQKRGPAEVVQKTA	1880	13.274	5.012	0.000	0.000	0.000	0.000	0.000
TNNQSSYPMPAVQKTA	1881	13.125	3.693	0.000	0.000	0.000	0.000	0.001
TNNQSRNPAEVVQKTA	1882	13.084	2.666	0.019	0.000	0.000	0.222	0.004
TNNQSNVRAEVVQKTA	1883	13.033	4.034	0.000	0.000	0.000	0.000	0.005
TNNQSKLLAEVVQKTA	1884	13.004	3.829	0.000	0.000	0.000	0.000	0.000
TNNQSYMIAEVVQKTA	1885	12.960	2.494	0.000	0.000	0.000	0.000	0.000
TNNQSSYPEPKVQKTA	1886	12.959	5.548	0.000	0.000	0.000	0.646	0.000
TNNQSSYPAESMLKTA	1887	12.928	2.621	0.000	0.000	0.002	0.144	0.001
TNNQSSYPAEILRKTA	1888	12.921	5.048	0.000	0.000	0.000	0.000	0.001
TNNQSFMVAEVVQKTA	1889	12.921	1.794	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAEKLVKTA	1890	12.911	4.402	0.005	0.000	0.000	0.168	0.004
TNNQSSYPAHSGQKTA	1891	12.828	3.582	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAENRGKTA	1892	12.790	4.245	0.000	0.000	0.023	0.000	0.003
TNNQSHLFAEVVQKTA	1893	12.741	2.798	0.000	0.000	0.000	0.000	0.000
TNNQSMTYAEVVQKTA	1894	12.738	0.925	0.000	0.000	0.000	0.000	0.000
TNNQSGRQAEVVQKTA	1895	12.598	2.184	0.000	0.000	0.000	0.000	2.796
IAPQSSYPAEVVQKTA	1896	12.595	0.000	0.000	0.000	0.000	0.000	0.000
TNNQSSYPAEKTGKTA	1897	12.539	4.257	0.000	0.000	0.000	0.000	0.000
TNNQSWALAEVVQKTA	1898	12.529	1.420	0.000	0.000	0.000	0.000	0.000
TNNQSSAPLEVVQKTA	1899	12.517	3.246	0.000	0.000	0.000	0.000	0.000
TNNQSFVKAEVVQKTA	1900	12.515	5.062	0.078	0.000	0.005	0.000	0.000
TNNQSSYPAEVGKKTA	1901	12.494	2.992	0.000	0.000	0.000	0.000	0.000

[1714]

TNNQSSYPA EVRKLTA	1902	12.48 6	3.95 2	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQYQSPA EVVQKTA	1903	12.28 8	2.13 4	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA ERQNKTA	1904	12.28 3	7.02 7	0.008	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYTA ESMLKTA	1905	12.25 0	0.89 1	0.000	0.00 0	0.00 0	0.000	0.00 0
TNLTSSYPA EVVQKTA	1906	12.09 3	3.00 6	0.000	0.00 0	0.00 0	0.000	1.81 8
TNNQSTVK AEVVQKTA	1907	12.08 5	3.35 7	0.009	0.00 0	0.00 0	0.000	0.00 0
TNNQSSPHT EVVQKTA	1908	11.97 4	1.81 1	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA VVMQKTA	1909	11.96 4	1.68 3	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSPGR EVVQKTA	1910	11.96 1	1.94 5	0.000	0.00 0	0.00 0	0.000	0.00 0
TIQTSSYPAE VVQKTA	1911	11.94 3	1.27 3	0.055	0.00 0	0.03 3	0.000	0.00 0
TNNQSSYPS HDVQKTA	1912	11.93 2	2.33 8	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSTIFE VVQKTA	1913	11.92 4	0.27 7	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSGKR AEVVQKTA	1914	11.86 9	1.71 0	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA EVSAKTA	1915	11.75 5	2.94 2	3.684	0.00 0	0.00 0	0.000	0.54 9
TNNQSTAPA EVVQKTA	1916	11.70 1	2.84 8	0.020	0.00 0	0.00 0	0.000	1.55 2
TNNQSVRN AEVVQKTA	1917	11.70 0	5.29 9	0.004	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA VNVQKTA	1918	11.68 3	2.85 5	0.000	0.00 0	0.00 3	0.624	0.00 0
TNNQSVFRA EVVQKTA	1919	11.63 6	3.42 3	0.000	0.00 0	0.01 0	0.000	0.00 0
TNNQSSYPA RDLQKTA	1920	11.63 5	2.64 9	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNRVTYPA EVVQKTA	1921	11.60 4	4.16 2	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYPA GASQKTA	1922	11.58 7	2.69 2	0.000	0.00 0	0.00 0	0.000	0.00 1
TNNQRLTPA EVVQKTA	1923	11.56 3	0.93 9	0.000	0.00 0	0.00 0	0.000	0.00 0
TNNQSKMF AEVVQKTA	1924	11.54 8	5.87 8	0.004	0.00 0	0.00 0	0.000	0.00 0
TNNQSSYTT MVVQKTA	1925	11.54 5	2.02 2	0.000	0.00 0	0.00 0	0.000	0.00 1
TNNQHSPPA	1926	11.52	1.60	0.090	0.00	0.00	0.000	0.00

EVVQKTA		4	2		0	0		1
TNNQSSYPA DHQQKTA	1927	11.518	2.922	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVVKNHA	1928	11.518	2.302	0.000	0.000	0.000	0.000	0.000
TNNQSKGIA EVVQKTA	1929	11.494	2.897	0.000	0.000	0.000	0.000	0.000
TNNPFVYPA EVVQKTA	1930	11.477	0.766	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA EVNQRTA	1931	11.467	5.421	0.000	0.000	0.000	0.000	0.017
TNNQSQAR AEVVQKTA	1932	11.462	3.052	0.000	1.765	0.000	0.000	0.000
TNAMRSYP AEVVQKTA	1933	11.435	0.717	0.297	0.000	0.000	0.000	0.008
TNNQSSYPA EMRTKTA	1934	11.414	6.798	0.000	0.000	0.000	0.000	0.000
TNNQSSYPM GTVQKTA	1935	11.385	1.645	0.000	0.000	0.000	0.000	0.000
TNNQSKLLA EVVQKPA	1936	11.342	2.915	0.000	0.000	0.000	0.000	0.000
TNNGHSYPA EVVQKTA	1937	11.323	1.066	0.000	0.000	0.123	0.000	0.006
TNNQSSYPA EKASKTA	1938	11.283	3.164	0.000	0.000	0.005	0.000	0.000
TNNQSYLPA EVVQKTA	1939	11.281	3.598	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA TASQKTA	1940	11.239	4.205	0.000	0.000	0.006	0.104	0.000
TNNQSSYPA EVGKSTA	1941	11.236	5.856	0.000	0.000	0.000	0.000	0.000
TNNQSSFNV EVVQKTA	1942	11.231	1.252	0.000	0.000	0.000	0.000	0.008
TNNQSSRYQ EVVQKTA	1943	11.191	2.650	0.000	0.000	0.000	0.000	0.000
TNNQQDLP AEVVQKTA	1944	11.184	0.000	0.000	0.000	0.000	0.000	0.000
TTSQSSYPA EVVQKTA	1945	11.183	1.308	0.000	2.070	0.000	0.000	0.000
TNNQSMGL AEVVQKTA	1946	11.182	3.666	0.000	0.000	0.000	0.000	1.522
TNNQSSILA EVVQKTA	1947	11.168	4.373	0.011	0.000	0.000	0.000	0.003
TNNQSSYPA RESQKTA	1948	11.167	5.693	0.000	0.000	0.000	0.000	0.002
TFFKSSYPA EVVQKTA	1949	11.161	6.543	0.000	0.000	0.000	0.000	0.000
TNNQSYLQ AEVVQKTA	1950	11.149	5.709	0.000	0.000	0.000	0.000	0.000

[1715]

[1716]

TNNQSSYPLKSVQKTA	1951	11.144	2.721	0.000	3.887	1.099	0.000	0.000
TNNQVSNPAEVVQKPA	1952	11.144	0.756	0.000	0.000	0.000	0.000	0.000
TNNQSSYPASSVKKTA	1953	11.121	2.694	0.000	0.000	0.000	0.000	0.000
SRSQSSYPAEVVQKTA	1954	11.116	3.092	0.000	0.000	0.000	0.000	0.000
TNNQSSYPATANQKTA	1955	11.110	1.662	0.006	0.000	0.000	0.000	0.008
TNNQRVSPAEVVQKTA	1956	11.107	2.954	0.000	0.491	0.000	0.000	0.006
TNNQSWTLAEVVQKTA	1957	11.074	5.721	0.000	0.000	0.000	0.000	0.000
TNNQSSTSQEVVQKTA	1958	11.070	5.925	0.000	0.000	0.000	0.000	0.004
TNNQSSTWSEVVQKTA	1959	11.024	4.711	0.104	0.000	0.000	0.000	1.778
TNNQSFRLAEVVQKTA	1960	11.023	3.075	0.000	0.000	0.000	0.000	0.000
TNNQSSYQREVVQKTA	1961	11.009	1.955	0.000	0.000	0.000	0.000	0.006
TNNQSSYPAEMTRKTA	1962	11.004	2.855	0.000	0.000	0.000	0.000	0.000
TNNQSSYPVHNVQKTA	1963	10.987	2.473	0.000	0.000	0.000	0.000	0.000
TNNQSSLHREVVQKTA	1964	10.972	2.605	0.000	0.000	0.000	0.000	0.000
TNNQSSFTREVVQKTA	1965	10.958	2.447	0.000	0.000	0.000	0.000	0.000
TNNQSSRVNEVVQKTA	1966	10.934	2.861	0.000	0.000	0.000	0.000	0.000
TNNQSSYPRGTVQKTA	1967	10.909	2.381	0.000	0.000	0.499	1.353	0.021
TPRPSSYPAEVVQKTA	1968	10.907	2.676	0.000	14.902	0.000	0.000	0.011
TNNQSRTYAEVVQKTA	1969	10.896	3.655	0.000	0.000	0.000	0.000	0.000
TNNQSSYTALVTQKTA	1970	10.889	7.450	0.000	0.000	0.000	0.000	0.000
TNNQSFNLAEVVQKTA	1971	10.863	5.748	0.036	0.000	0.000	0.000	0.002
TNWQSSYPAEVVQKTA	1972	10.804	3.863	0.000	0.000	0.000	0.000	0.000
TLQTSSYPAEVVQKTA	1973	10.800	3.298	0.012	77.694	374.6	0.000	0.000
TNNQSSYPAERLQKTA	1974	10.798	4.850	0.000	0.000	0.000	0.000	0.000
TNNQSSMLP	1975	10.79	3.73	0.001	0.00	0.00	0.000	0.00

[1717]

EVVQKTA		2	5		0	0		1
TNNQSSYPESGVQKTA	1976	10.772	2.301	0.000	0.000	0.000	0.000	0.000
TNNQSSYPSTNVQKTA	1977	10.723	3.581	0.000	0.000	0.000	0.000	0.005
TNNQSSYPFSQVQKTA	1978	10.655	4.068	0.000	0.000	0.000	0.000	0.000
TNNQSGQLAEVVQKTA	1979	10.620	2.718	0.000	0.000	0.000	0.000	0.003
TNNQSSRWQEVVQKTA	1980	10.619	0.039	0.000	0.000	0.000	0.000	0.000
TNNQSSSPAEVVQKPA	1981	10.609	4.216	0.000	0.000	0.000	0.000	0.299
TNNQSSIALEVVQKTA	1982	10.601	4.298	0.000	0.000	0.000	0.000	0.000
TNNQSSARSEVVQKTA	1983	10.585	4.080	0.009	0.000	0.000	0.000	0.003
TNNQSSSVREVVQKTA	1984	10.563	2.732	0.000	0.000	0.000	0.000	0.000
TNNQSMSHAEVVQKTA	1985	10.559	2.887	0.000	0.000	0.000	0.000	0.000
TNNQSSKTQEVVQKTA	1986	10.559	1.692	0.000	1.914	0.000	0.000	0.002
TNNQSSIGTEVVQKTA	1987	10.505	2.615	0.000	0.000	0.000	0.000	0.009
TNNQSSYPASAMQKTA	1988	10.505	3.785	0.001	0.000	0.000	0.000	0.000
TNNQSSYPAQHTQKTA	1989	10.502	3.440	0.000	0.000	0.000	0.000	0.000
TNNQSSYPPKQVQKTA	1990	10.501	1.690	0.000	0.000	0.000	0.000	0.006
TNSKNSYPAEVVQKTA	1991	10.499	2.291	0.033	0.000	0.000	0.000	0.000
TNNQRSMPAEVVQKTA	1992	10.485	3.812	0.050	0.000	0.000	0.000	0.000
TNNQSSSNLEVVQKTA	1993	10.475	3.045	0.000	0.000	0.000	0.000	0.000
TNNQVLYPAEVVQKTA	1994	10.467	2.334	0.005	0.000	0.016	0.000	0.001
TNNQSIHPAEVVQKTA	1995	10.463	2.236	0.000	0.000	0.000	0.248	0.003
TNNQSYNMAEVVQKTA	1996	10.429	5.146	0.000	0.000	0.000	0.000	0.000
TNNQSSYTASSAQKTA	1997	10.427	3.158	0.000	0.000	0.000	0.000	0.000
TNNQSSNGVEVVQKTA	1998	10.415	2.651	0.000	0.000	0.000	0.000	0.000
TNNQSSYPDRLVQKTA	1999	10.409	2.906	0.001	0.000	0.000	0.000	0.000

[1718]

TNNQSSLLA EVVQKTA	2000	10.367	3.838	0.000	23.928	0.011	2.016	0.090
TNNQSPTAA EVVQKTA	2001	10.307	1.949	0.000	0.000	0.000	0.000	0.001
TNNQSSQLA EVVQKTA	2002	10.294	3.792	0.008	0.000	0.005	0.000	0.000
TNNQSSYPT TEVQKTA	2003	10.277	5.539	0.000	0.000	0.000	0.000	0.004
TNNQSSYPA EASQKTA	2004	10.233	3.979	0.007	0.000	0.000	0.000	0.000
TNNQSWVL AEVVQKTA	2005	10.223	2.827	0.000	0.000	0.000	0.000	0.000
TNNQSSIVR EVVQKTA	2006	10.214	4.402	0.000	0.000	0.000	0.178	0.002
TNNQWTRP AEVVQKTA	2007	10.195	1.655	0.000	0.000	0.000	0.000	0.000
TNNQSSYER VVVQKTA	2008	10.189	0.652	0.000	0.000	0.000	0.000	0.000
TNNQSSVRS EVVQKTA	2009	10.180	3.489	0.001	0.000	0.000	0.000	0.000
TNNQSSELR EVVQKTA	2010	10.174	2.995	0.000	0.000	0.015	0.000	0.000
TNNQSSYVK SVVQKTA	2011	10.171	2.873	0.000	0.000	0.000	0.000	0.000
TNNQSSYPK LQVQKTA	2012	10.161	3.023	0.000	0.000	0.000	0.000	0.000
TNNQSSYDV MVVQKTA	2013	10.132	6.923	0.000	0.000	0.000	0.000	0.000
TNSNVSYPA EVVQKTA	2014	10.114	2.151	0.000	0.000	0.008	0.000	0.000
TNNQWSMP AEVVQKTA	2015	10.109	2.452	0.026	0.000	0.000	0.000	0.000
TNNQSSSQL EVVQKTA	2016	10.102	3.295	0.000	0.000	0.000	0.309	1.567
TNNQSTAN AEVVQKTA	2017	10.084	5.631	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA TGSQKTA	2018	10.069	5.184	0.000	0.000	0.000	0.000	0.001
TNNQSSPSL EVVQKTA	2019	10.046	7.144	0.000	0.000	0.000	0.000	0.002
TNNAYTYP AEVVQKTA	2020	10.027	0.424	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA SVVKKTA	2021	10.026	0.746	7.839	0.000	0.000	0.000	4.347
TYAKSSYPA EVVQKTA	2022	10.011	1.568	0.000	0.000	0.000	0.000	0.000
TNNQSSYPA VGGQKTA	2023	9.968	2.899	0.000	0.000	0.000	0.000	0.000

[1719] 这些数据表明,在两种成熟方法后,与相应未成熟衣壳变体相比,在NHP中生成具有显著增强的CNS、心脏、肌肉和肝趋向性的具有环VIII修饰的成熟TTN-002衣壳变体(AAV5

衣壳变体),其已在小鼠、大鼠和/或NHP的脑中展现相对于AAV5和/或AAV9的显著倍数富集。此外,数种所生成的成熟变体在NHP和小鼠中证实交叉物种CNS趋向性。

[1720] 实施例5.TTN-002AAV衣壳变体在不同灵长类动物物种中的评估

[1721] 此实例评估TTN-002(SEQ ID NO:982(氨基酸)和984(DNA),包含SEQ ID NO:943)衣壳变体在两种不同灵长类动物物种狨猴(普通狨)和非洲绿猴(西非绿猴)中的趋向性和交叉物种兼容性,如与其在实施例1和2中所提供的食蟹猕猴中的趋向性相比。TTN-002衣壳变体的氨基酸和DNA序列分别提供于例如表4和表5中。

[1722] 为了研究非洲绿猴中的趋向性,将包含TTN-002衣壳变体的AAV颗粒或处于突触蛋白启动子的控制下的AAV5对照以2E13个vg/kg的剂量静脉内注射到非洲绿猴(n=2,3-12岁)中。在出生14天后,收集NHP的脑和组织(肝、DRG、股四头肌和心脏)且提取RNA。在RNA回收和RT-PCR扩增之后,执行系统性NGS富集分析以计算相对于AAV5野生型对照的倍数富集比率。

[1723] 为了研究狨猴中的趋向性,将包含TTN-002衣壳变体的AAV颗粒或AAV5对照以2E13个vg/kg的剂量静脉内注射到狨猴(n=2,>10月龄)中。在出生28天后,收集NHP的脑和组织(肝、股四头肌和心脏)且提取RNA。在RNA回收和RT-PCR扩增之后,执行系统性NGS富集分析以计算相对于AAV5野生型对照的倍数富集比率。

[1724] 如表21(非洲绿猴)和表22(狨猴)中所提供,TTN-002衣壳变体在不同灵长类动物物种中证实增加的CNS趋向性。TTN-002衣壳变体在食蟹猕猴的脑中证实相对于AAV5的64.9倍表达增加(表9,实施例1),在非洲绿猴的脑中证实相对于AAV5的7.5倍表达增加,且在狨猴的脑中证实相对于AAV5的40.4倍表达增加。此外,TTN-002还导致大鼠的脑中的表达增加(表9,实施例1),这证实表达相对于AAV5的平均倍数变化为41.1。

[1725] 表21.非洲绿猴中TTN-002的NGS倍数富集

[1726]

序列	SEQ ID NO:	相对于 AAV9 的倍数富集					
		脑	DRG	心脏	肝 DNA	肝 RNA	肌肉
YPAEVVQK	943	7.462	3.9674	0.3592	0.0543	0.0227	0.6982

[1727] 表22.狨猴中TTN-002的NGS倍数富集

[1728]

序列	SEQ ID NO:	相对于 AAV9 的倍数富集				
		脑	心脏	肝 DNA	肝 RNA	肌肉
YPAEVVQK	943	40.381	3.1981	0.0321	0.0328	2.0631

[1729] 总之,这些数据表明,AAV5衣壳变体TTN-002在三种不同灵长类动物物种和大鼠中的CNS中证实相对于AAV5对照的CNS趋向性增加,从而提供强交叉物种能力的证据。

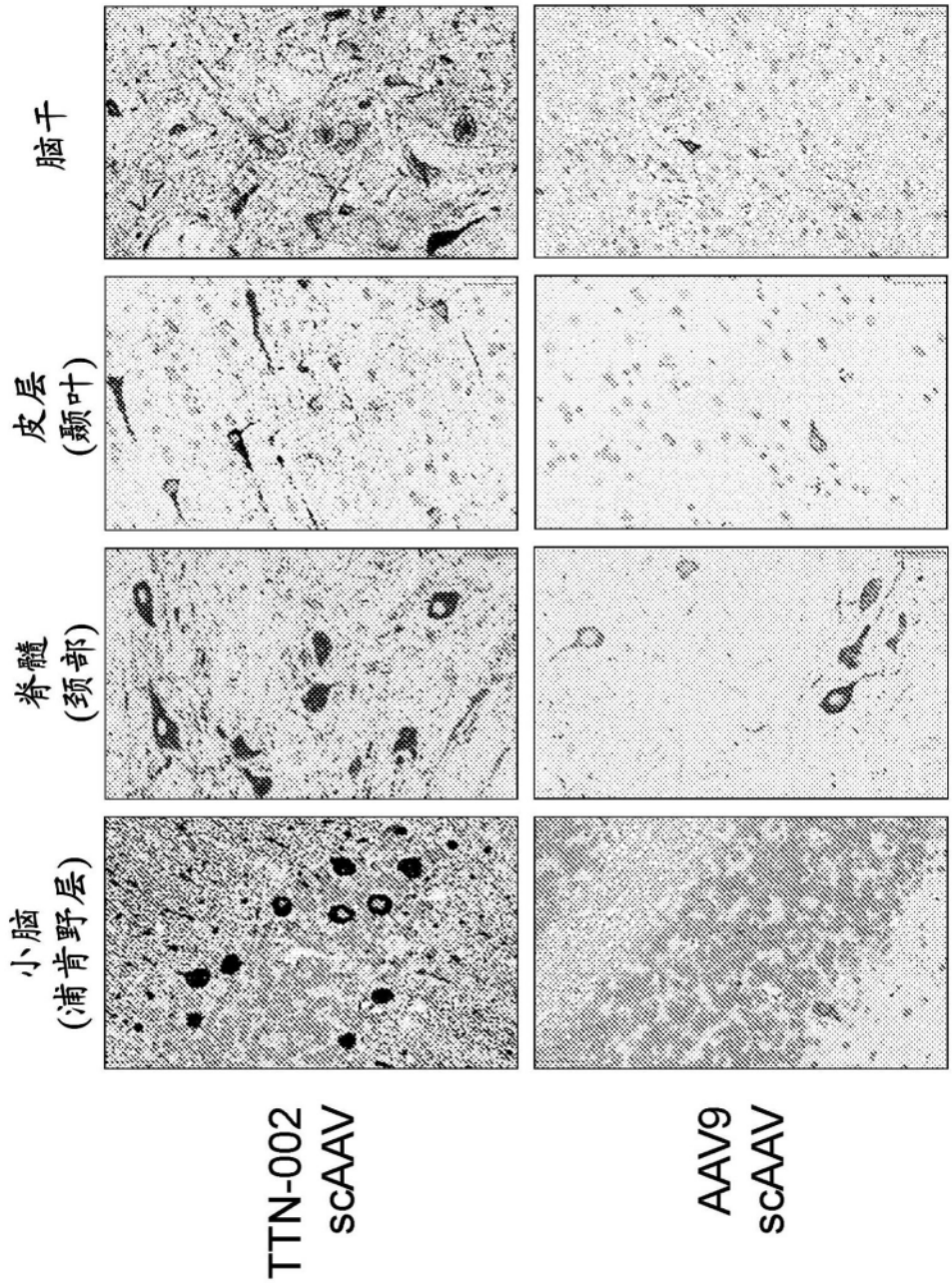

图1A

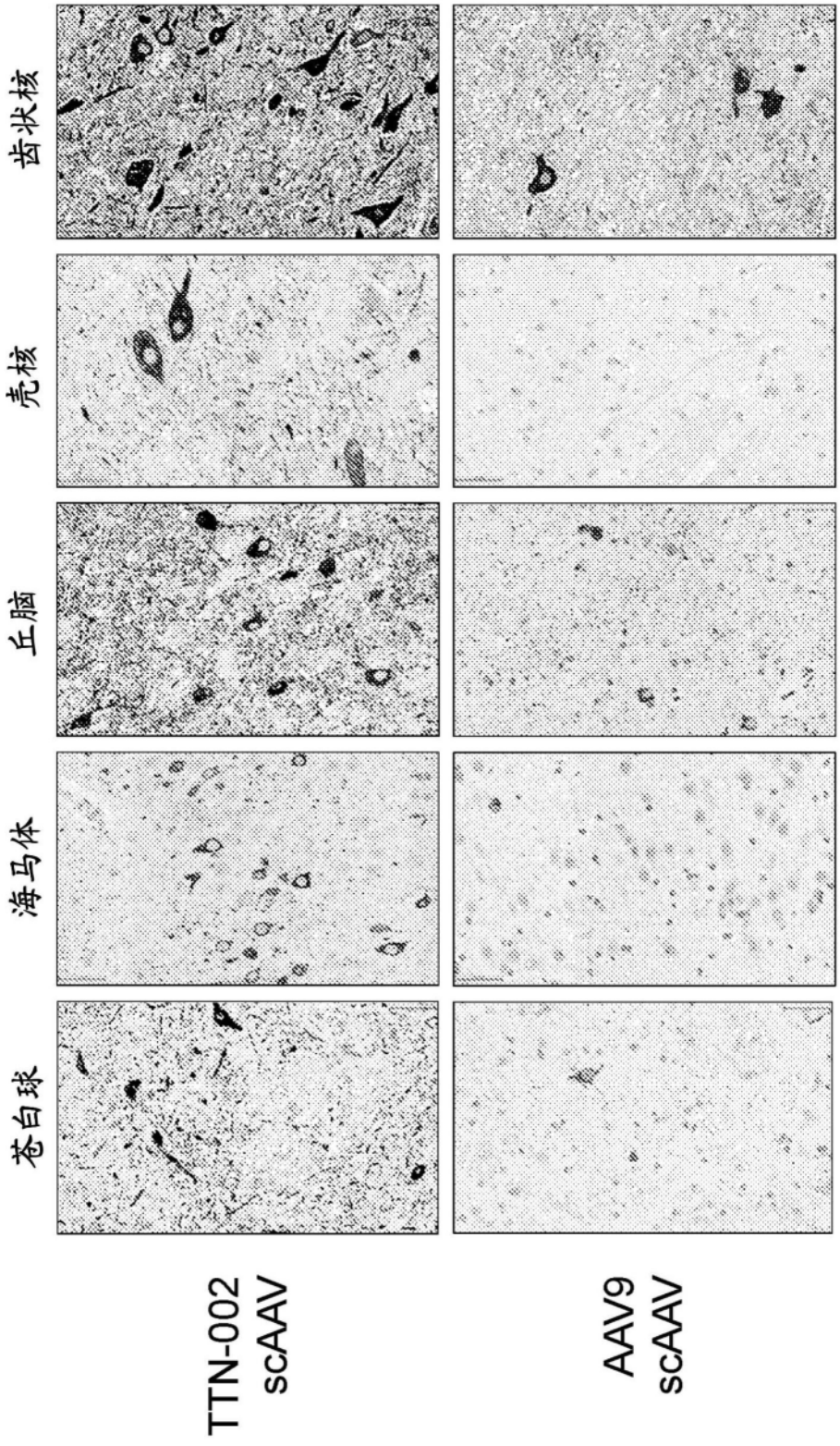

图1B

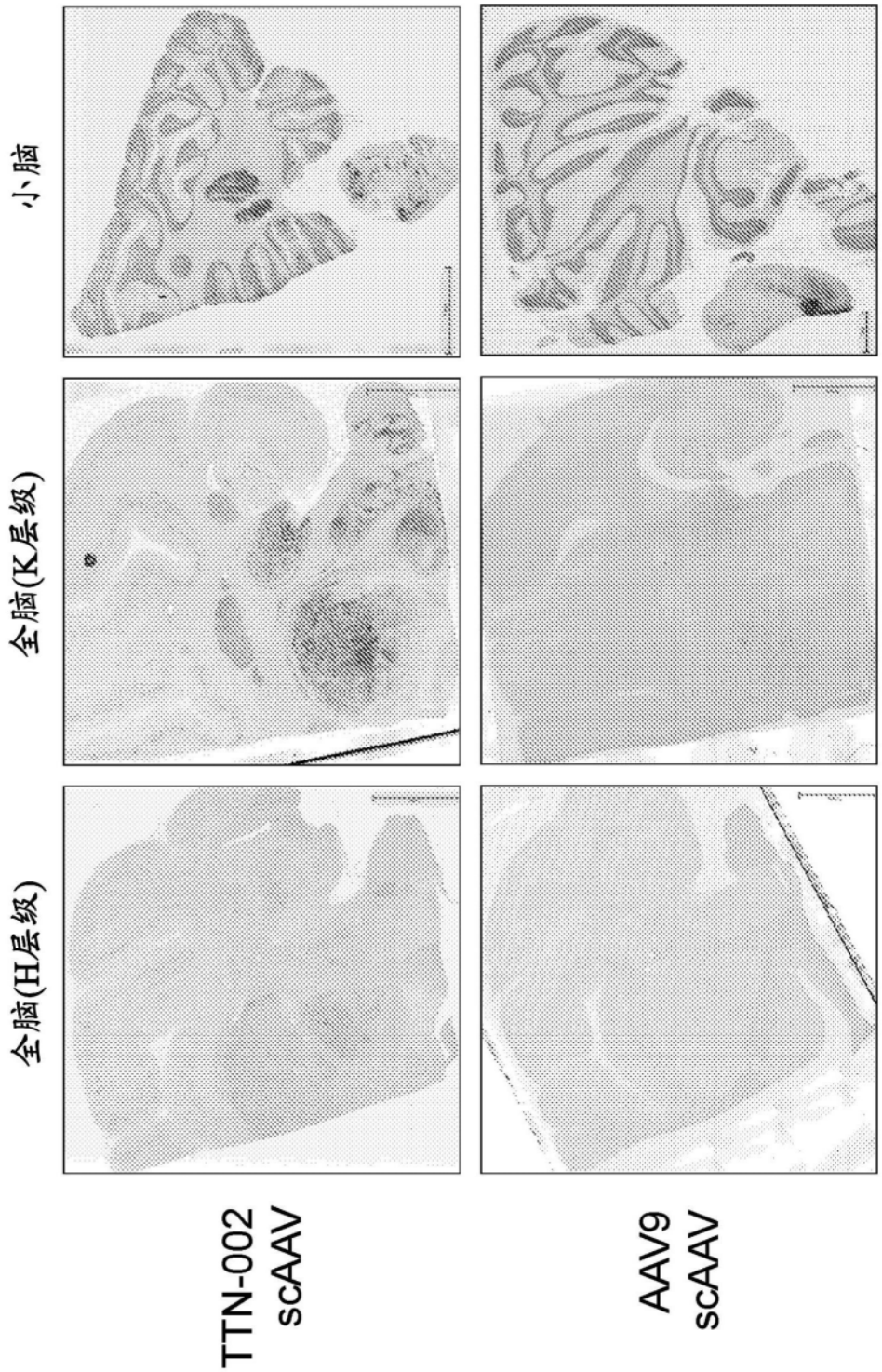

图1C

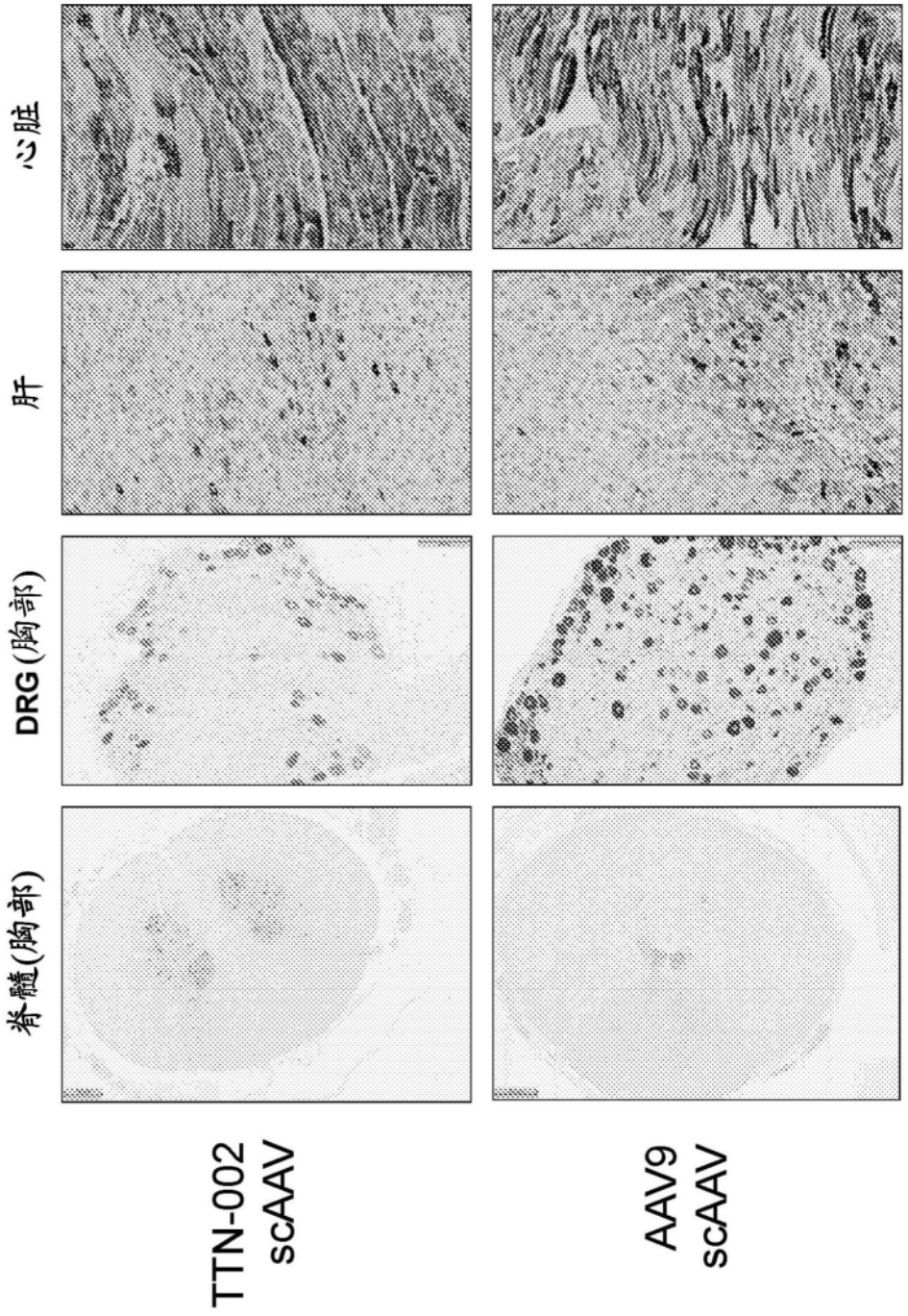

图1D